Affymetrix® Data Exchange Console 2.0 User's Guide

P/N 702571 Rev. 3

For research use only.
Not for use in diagnostic procedures.

Trademarks

Affymetrix®, , GeneChip®, HuSNP®, GenFlex®, Flying Objective™, CustomExpress®, CustomSeq®, NetAffx®, Tools To Take You As Far As Your Vision®, The Way Ahead™, Powered by Affymetrix™, GeneChip-compatible™, Command Console®, and GeneTitan™ are trademarks of Affymetrix, Inc.

All other trademarks are the property of their respective owners.

Limited License Notice

Limited License. Subject to the Affymetrix terms and conditions that govern your use of Affymetrix products, Affymetrix grants you a non-exclusive, non-transferable, non-sublicensable license to use this Affymetrix product only in accordance with the manual and written instructions provided by Affymetrix. You understand and agree that except as expressly set forth in the Affymetrix terms and conditions, that no right or license to any patent or other intellectual property owned or licensable by Affymetrix is conveyed or implied by this Affymetrix product. In particular, no right or license is conveyed or implied to use this Affymetrix product in combination with a product not provided, licensed or specifically recommended by Affymetrix for such use.

Patents

Scanner products may be covered by one or more of the following patents: U.S. Patent Nos. 5,578,832; 5,631,734; 5,834,758; 5,936,324; 5,981,956; 6,025,601; 6,141,096; 6,171,793; 6,185,030; 6,201,639; 6,207,960; 6,218,803; 6,225,625; 6,252,236; 6,335,824; 6,403,320; 6,407,858; 6,472,671; 6,490,533; 6,650,411; 6,643,015; 6,813,567; and other U.S. or foreign patents.

Software products may be covered by one or more of the following patents: U.S. Patent Nos. 5,733,729; 5,795,716; 5,974,164; 6,066,454; 6,090,555; 6,185,561; 6,188,783; 6,223,127; 6,228,593; 6,229,911; 6,242,180; 6,308,170; 6,361,937; 6,420,108; 6,484,183; 6,505,125; 6510,391; 6,532,462; 6,546,340; 6,687,692; 6,607,887; 7,062,092 and other U.S. or foreign patents.

Fluidics stations Products may be covered by U.S. Patent No. 6,114,122; 6,287,850; 6,391,623; 6,422,249; and other U.S. or foreign patents.

AutoLoader products may be covered by one or more of the following patents: U.S. Patent Nos. 6,511,277; 6,604,902; 6,705,754; 7,108,472; and other U.S. or foreign patents.

Copyright

© 2008 Affymetrix, Inc. All Rights Reserved.

TABLE OF CONTENTS

Chapter 1 Welcome to the Data Exchange Console 1

Data Migration 1
Conventions Used in This Guide 2
Resources 3

Chapter 2 Introduction to Data Migration. 5

Starting the Data Exchange Console 5
GCOS/AGCC File Differences 6
File Format Impacts 9
Import Conflicts and Errors 9
Export Conflicts 10
Importing HT Array Plate Scanner Files 11

Chapter 3 DEC Import 13

Selecting Import Method 13
Importing DTT Files 15
Importing From GCOS Connect 22
Importing MAS/GCOS Flat Files 27
Importing HT Array Plate Scanner Data 31

Chapter 4 DEC Export. 37

Choosing How to Select Files for Export 37
Using the Filter 38
Selecting Sample Files for Export 40
Selecting a Location 41
Reviewing Data and Fixing Conflicts 42
Export Status 45

Chapter 5 Impacts on Data 47

Impacts Caused by Migrating from GCOS to AGCC 47
Impacts Caused by Migrating from AGCC to GCOS 50

Chapter 1

WELCOME TO THE DATA EXCHANGE CONSOLE

This manual explains the process of migrating data from various versions of GeneChip® Operating Software (GCOS) software into Affymetrix® GeneChip® Command Console® (AGCC) software and back. The migration is performed using the GCOS Data Transfer Tool (DTT) and the Affymetrix Data Exchange Console (DEC).

The manual includes:

- A discussion of the differences between GCOS and AGCC (see Chapter 2 *Introduction to Data Migration* on page 5).
- Instructions on using the Data Exchange Console to import data into and export data from AGCC.

  See:

  - Chapter 3 *DEC Import* on page 13
  - Chapter 4 *DEC Export* on page 37
- Descriptions of the possible impacts on data that may result (see Chapter 5 *Impacts on Data* on page 47).

The DEC Installation installs:

The Data Exchange Console install package installs:

- Data Exchange Console (DEC)
- Affymetrix Command Console Library Importer Tool
- Affymetrix Command Console Reconnector

The installation of DEC and the other tools is described in the *AGCC 2.0 Installation Instructions*.

The use of the Library Importer Tool is described in the *AGCC 2.0 Installation Instructions*.

The use of the Reconnector is described in the *Affymetrix Command Console Reconnector 2.0 User's Guide*.

See *Affymetrix® Data Transfer Tool User's Guide Version 1.1* for information on using the DTT.

Data Migration

You can perform a data migration:

- *From GCOS to AGCC*

  This allows you to transfer previously generated data into AGCC for organization and analysis.
- *From AGCC to GCOS*

  This allows you to transfer data from AGCC to a computer with GCOS installed if you are dealing with a user who has not upgraded yet.

From GCOS to AGCC

You can migrate data between GCOS systems and AGCC systems using several different methods:

- You can use the Data Transfer Tool (DTT) to export data from GCOS in one of the DTT file formats, and then import the DTT files into AGCC using DEC. This enables you to transfer experiment and sample information, as well as the Data files (DAT, CEL, and CHP).
- You can connect directly between the GCOS system and the AGCC system using DTT. This will also enable you to transfer experiment and sample information from the GCOS process database.
- You can import the flat data files (in MAS 5.x or GCOS 1.x formats) into AGCC using DEC. This will not import experiment and sample information.

For more information, see Chapter 3 *DEC Import* on page 13.

From AGCC to GCOS

You can migrate data from AGCC to GCOS by using DEC to export AGCC data into DTT format, which can then be imported into a GCOS system. This is useful when dealing with users who have not updated to AGCC yet.

For more information, see Chapter 4 *DEC Export* on page 37.

NOTE: CHP files created with the new expression algorithms (e.g., PLIER, RMA, iterPLIER) and genotyping algorithms (e.g., BRLMM_P, Birdseed) cannot be loaded into GCOS and will not be included in the DTT file.

NOTE: HT data from AGCC cannot be imported back into GCOS.

Conventions Used in This Guide

This manual provides a detailed outline for all tasks associated with the Data Exchange Console. Various conventions are used throughout the manual to help illustrate the procedures described. Explanations of these conventions are provided below.

Steps

Instructions for procedures are written in a step format. Immediately following the step number is the action to be performed. Following the response additional information pertaining to the step may be found and is presented in paragraph format. For example:

1. Click **Yes** to continue.

   The Delete task proceeds.
   In the lower right pane the status is displayed.

   To view more information pertaining to the delete task, right-click **Delete** and select **View Task Log** from the shortcut menu that appears.

Font Styles

Bold fonts indicate names of commands, buttons, options or titles within a dialog box. When asked to enter specific information, such input appears in italics within the procedure being outlined.
For example:

1. Click the **Find** button or select **Edit** → **Find** from the menu bar.

   The Find dialog box appears.

2. Enter *AFFX-BioB-5_at* in the **Find what** box, then click **Find Next** to view the first search result.

3. Continue to click **Find Next** to view each successive search result.

Screen Captures

The steps outlining procedures are frequently supplemented with screen captures to further illustrate the instructions given. The screen captures depicted in this manual may not exactly match the windows displayed on your screen.

Additional Comments

TIP: Information presented in Tips provide helpful advice or shortcuts for completing a task.

NOTE: The Note format presents important information pertaining to the text or procedure being outlined.

IMPORTANT: The Important format presents important information that may affect the accuracy of your results.

CAUTION: Caution notes advise you that the consequence(s) of an action may be irreversible and/or result in lost data.

WARNING: Warnings alert you to situations where physical harm to person or damage to hardware is possible.

Resources

Online Documentation

This manual is available in Adobe Acrobat format (as *.pdf files) in the download package for AGCC and is readable with the Adobe® Acrobat Reader® software, available at no charge from Adobe at http://www.adobe.com.

The content of this manual is available in context-sensitive online help, accessible by clicking the Help buttons in the Data Exchange Console.

Technical Support

Affymetrix provides technical support to all licensed users via phone or E-mail. To contact Affymetrix® Technical Support:

Affymetrix, Inc.

34200 Central Expressway

Santa Clara, CA 95051 USA

E-mail: support@affymetrix.com

Tel: 1-888-362-2447 (1-888-DNA-CHIP)

Fax: 1-408-731-5441

Affymetrix UK Ltd.

Voyager, Mercury Park

Wycombe Lane, Wooburn Green

High Wycombe HP10 0HH

United Kingdom

UK and Others Tel: +44 (0) 1628 552550

France Tel: 0800919505

Germany Tel: 01803001334

E-mail: supporteurope@affymetrix.com

Tel: +44 (0) 1628 552550

Fax: +44 (0) 1628 552585

Affymetrix Japan, K. K.

Mita NN Bldg

16 Floor, 4-1-23 Shiba,

Minato-ku, Tokyo 108-0014

Japan

Tel: +81 (03) 5730-8222

Fax: +81 (03) 5730-8201

www.affymetrix.com

Chapter 2

INTRODUCTION TO DATA MIGRATION

The Data Exchange Console (DEC) for AGCC and the Data Transfer Tool (DTT) for GCOS allow you to migrate data:

- From GCOS to AGCC
- From AGCC to GCOS

GCOS and AGCC use different systems to organize the data and different formats for the data files themselves. These differences are a potential source of problems when you migrate data, and can have impacts on the data. For more information, see:

- *GCOS/AGCC File Differences* on page 6
- *File Format Impacts* on page 9

In addition, inconsistencies in the data or duplicated names in GCOS and AGCC can cause conflicts during data import. For more information, see *Import Conflicts and Errors* on page 9.

These types of inconsistencies can also cause conflicts when transferring data from AGCC to GCOS. For more information, see *Export Conflicts* on page 10.

Importing HT Array Plate Scanner data involves some special issues. See *Importing HT Array Plate Scanner Files* on page 11 for more information.

Starting the Data Exchange Console

To start the Data Exchange Console:

- Click the Microsoft® Windows® **Start** button and select **Programs** → **Affymetrix** → **Data Exchange Console**.

  The Welcome window opens (Figure 2.1).

The Welcome window opens when you first start DEC.

This is where you can choose:

- Import Data
- Export Data

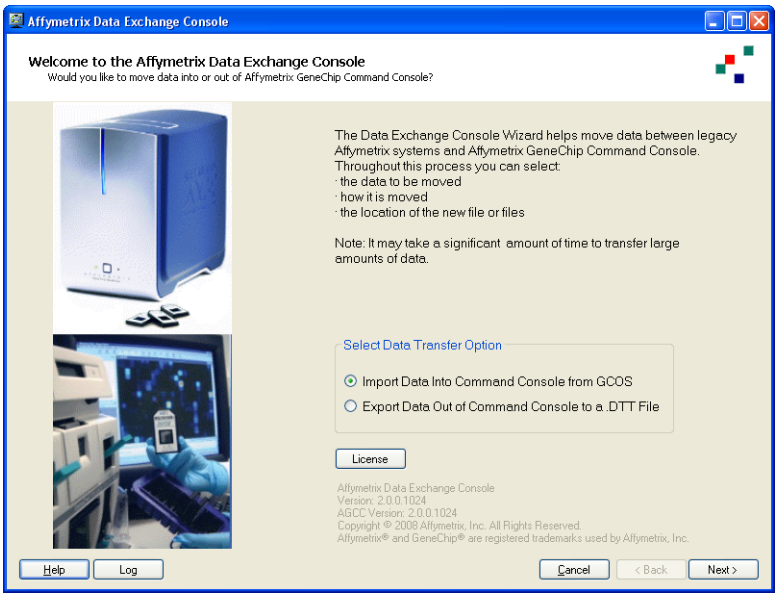
Affymetrix Data Exchange Console
Welcome to the Affymetrix Data Exchange Console
Would you like to move data into or out of Affymetrix GeneChip Command Console?
The Data Exchange Console Wizard helps move data between legacy Affymetrix systems and Affymetrix GeneChip Command Console. Throughout this process you can select:
· the data to be moved
· how it is moved
· the location of the new file or files
Note: It may take a significant amount of time to transfer large amounts of data.
Select Data Transfer Option
Import Data Into Command Console from GCOS
Export Data Out of Command Console to a .DTT File
License
Affymetrix Data Exchange Console
Version: 2.0.0.1024
AGCC Version: 2.0.0.1024
Copyright © 2008 Affymetrix, Inc. All Rights Reserved.
Affymetrix® and GeneChip® are registered trademarks used by Affymetrix, Inc.
Help
Log
Cancel
< Back
Next >

Figure 2.1 Welcome window

Use the buttons at the bottom to:

Help	Open the Online help.
Log	Displays a text file with a record of the file imports.
Cancel	Close the dialog box.
Back	Return to the previous step (disabled in this window).
Next	Go to the next step.

To import data:

1. Select **Import Data into Command Console from GCOS**.
2. Click **Next**.
   The Select Import Method window opens (see page 13).

To export data:

1. Select **Export Data out of Command Console to a DTT file**.
2. Click **Next**.
   The Select Export Options window opens (see page 37).

GCOS/AGCC File Differences

There are some differences between the way files are organized and the data is stored in AGCC and GCOS. You need to understand these differences to understand some of the problems that may arise when importing or exporting data.

GCOS Files and Database

In GCOS, the sample and experiment information is kept in the process database as attributes of the sample and experiment.The information is assigned by the user during experiment registration.

The database includes information about the data files and about the fluidics and scanner operation.

The data generated during the experiment is in individual data files:

- Image (DAT) files
- Intensity (CEL) Files
- Probe Analysis (CHP) Files
- Report (RPT) Files

The format of these file types may have changed in different versions of GCOS. For more information, see *Impacts Caused by Migrating from GCOS to AGCC* on page 47.

The information and data are organized by:

- Project
- Sample
- Experiment

Sample and Experiment Templates are used to organize the attributes and their data characteristics; the templates can be selected for a sample or experiment during registration.

Templates can be edited to add or remove attributes or change their data characteristics.

The GCOS database also stores algorithm settings for the different types of analysis (expression, mapping, etc.). These settings are organized into usersets, which can also be copied and edited.

AGCC Files

When GCOS data is transferred into AGCC, the sample and experiment attributes and other information are placed in the following XML files:

- Sample files (ARR)
- Audit Information files (AUDIT)

The experiment data is stored in:

- DAT Files
- CEL files
- CHP files
- MSK files

These files contain the same type of data as their counterparts in GCOS, but the format is different and some data may not be included.

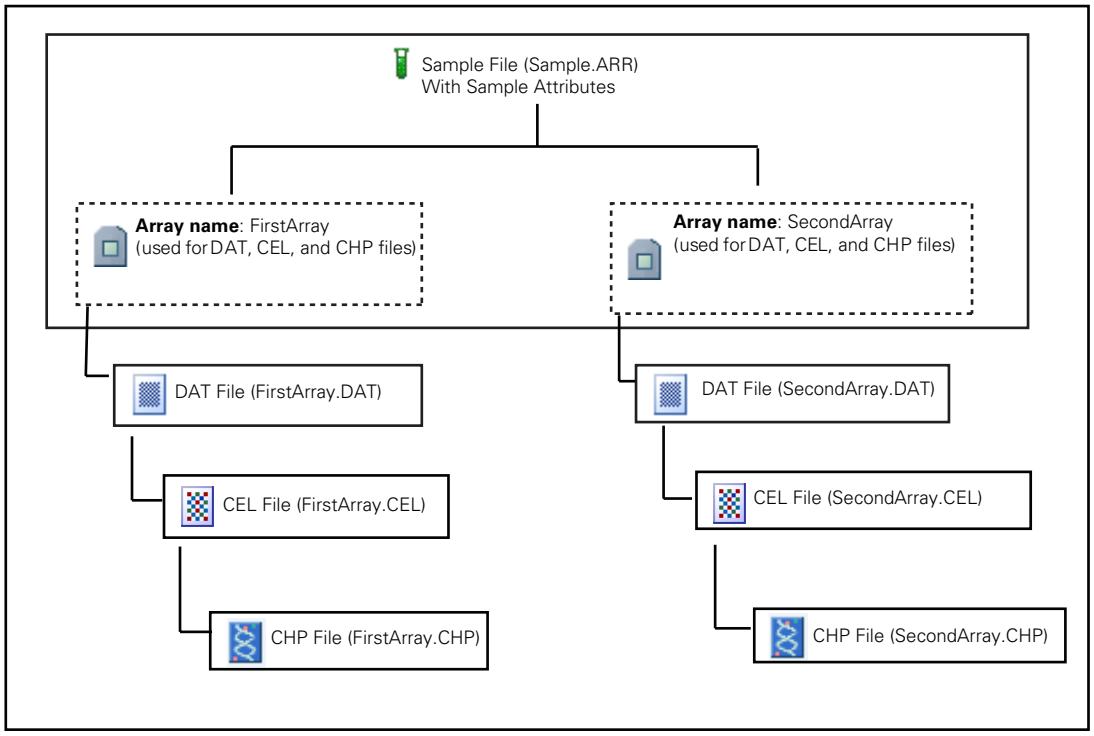

Figure 2.2 Sample File with multiple arrays

The files can be assigned to a Project during import, or you can create subfolders to preserve the GCOS project structure. These subfolders are added to their respective projects in AGCC.

When data is imported from GCOS, the userset information on analysis algorithm parameters is kept in Parameter Files (.PARAM). These files are not visible in the AGCC interface.

Template information is placed into a template file that can be edited through the AGCC Portal.

See Chapter 2 *Getting Started With AGCC on page 5* of the *AGCC User's Guide* for more information about the file types and data organization in AGCC.

Tracking Files with GUIDs

A GUID, or Globally Unique Identifier, is assigned to each file for tracking (Figure 2.3). GUIDs are numbers generated to track a file that will be unique to that file.

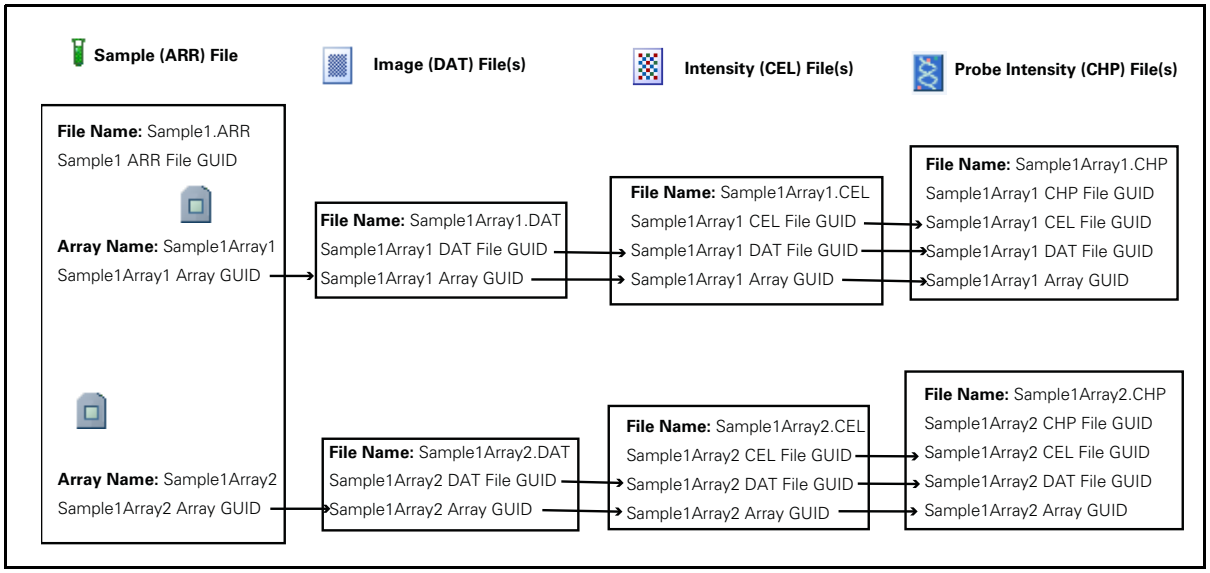

Figure 2.3 GUIDs in Sample and Data Files

During Sample registration the Sample file is assigned a Sample File GUID; in addition an Array GUID is provided for every array name entered.

Every data file (DAT, CEL, and CHP) generated for an array will contain the Array GUID for the array, as well as the GUIDs for each of its parent data files.

The GUIDs enable you to trace the lineage of any data file independent of the file name.

The parent-child relationships can be broken under certain circumstances when using DEC; see *Affymetrix Command Console Reconnector User's Guide* for information on repairing these broken links.

File Format Impacts

There are differences in the way AGCC stores data, as opposed to GCOS. These differences can cause impacts when you copy data from one to the other.

The impacts may vary, depending upon:

- Which way the transfer takes place (GCOS to AGCC or AGCC to GCOS)
- The software source for the data being imported (MAS 5.x and the different versions of GCOS)
- The type of file that is being transferred (DAT, CEL, and CHP)
- The array type (expression, reseq, mapping, etc.)
- The type of algorithm used for processing the data.
- Possibly the actual transfer file format used (DTT, CAB, MAS 5.x)

You need to be aware of the possible impacts before transferring data. You can learn more about file format impacts in Chapter 5 *Impacts on Data* on page 47.

Import Conflicts and Errors

Import conflicts occur when there is a discrepancy or duplication in the data to be imported. The discrepancy can be between the data to be imported and the data in AGCC already, or between two files in the set of data to be imported.

These conflicts can be detected by DEC and resolved by correcting the discrepancy or renaming the item to be imported that is causing the discrepancy.

These discrepancies can involve:

- *File Names*, below
- *Sample and Experiment Template Names* on page 10
- *Usersets* on page 10
- *Physical Array Name Conflicts* on page 10
- *Unresolvable Conflicts Notices* on page 10

For more information about resolving conflicts, see the relevant sections of Chapter 3 *DEC Import* on page 13.

An error is a conflict or other problem that cannot be resolved.

File Names

A file name conflict arises when:

- You are trying to import a file with the same name as a file that is already in AGCC.
- You are trying to import two different files with the same name.

Sample and Experiment Template Names

Template conflicts can be caused by:

- Discrepancies in the attributes in the template used in AGCC and the attributes in the templates used for the data being imported.
- Discrepancies between two different templates being imported into AGCC.
- Duplicate attribute names in a template.
- Discrepancies in the Data Types and control values for attributes in AGCC templates and templates to be imported.
- Sample and experiment templates with the same template name

Usersets

Userset conflicts occur when the imported userset has the same name as a userset already in Command Console but with different values or probe array types.

Physical Array Name Conflicts

Physical array name conflicts between imported data set and physical arrays already in AGCC are not considered.

Unresolvable Conflicts Notices

Typical unresolvable conflicts involve:

- A barcode is used by a sample file already in AGCC.

Export Conflicts

Certain conflicts can arise when migrating data from AGCC to GCOS.

Sample File User Attribute Names

Due to namespaces in user attributes, samples in AGCC can have multiple attributes with the same name. GCOS can not handle this situation so these attributes need to be renamed for export to GCOS.

Data File or Experiment Name conflicts

Within the data being exported, since AGCC can have many Sample and Data files with the same name, but in different folders, there can be name conflicts when exporting data into a DTT file.

Exported DAT and CEL Names Mismatch

DTT expects that a CEL file that is based on a DAT file will have the same root name as the DAT file. If the CEL file does not, then DTT will fail to import it. To get around this limitation, DEC will rename the DAT file to match the child CEL file and show a conflict resolution. The user may change the root name of the DAT and CEL file. A consequence of this limitation is that only one CEL file per DAT file may be successfully exported.

Different Samples With the Same Name

In AGCC, many samples can have the same name and have different attributes or belong to different projects. DEC will assume that every sample file is distinct regardless of the name. When two samples with the same name are added to the same DTT file, DEC shall rename the samples and allow the user to change the name through the conflict resolution dialog.

Importing HT Array Plate Scanner Files

HT array plates can be processed in GCOS. For each array on a plate, GCOS produces:

- 84 or 98 mini-DAT files
- 1 CEL file
- 1CHP file

When the GCOS HT Array Plate data is imported into AGCC using DEC, the following AGCC files for each array are created:

- 1 DAT file
- 1 CEL file
- 1 CHP file

The mini-DATs get consolidated during import, and the CEL and CHP files get imported in the AGCC format.

See Chapter 8 *Using the AGCC Viewer* in the *AGCC 2.0 User's Guide* for more information about how HT Array Plates Scanner files are set up.

Chapter 3

DEC IMPORT

You can import GCOS data into AGCC using several different methods. For any import workflow, you follow the same set of steps:

1. Select the import option you wish to use.
2. Select the file(s) to be imported.
3. Select experiments from the file (optional).
4. Select or create a project for the imported files.
5. Check for conflicts.
6. Resolve conflicts if necessary.
7. Import data.

Selecting Import Method

You can import data using the following methods (Figure 3.1):

- Importing DTT 1.1 Archive or Flat Files (see page 15)
- Using GCOS Connect (see page 22)
- Importing MAS 5.x, GCOS 1.x Flat Files (see page 27)
- Importing HT Array Flat Files (see page 31)

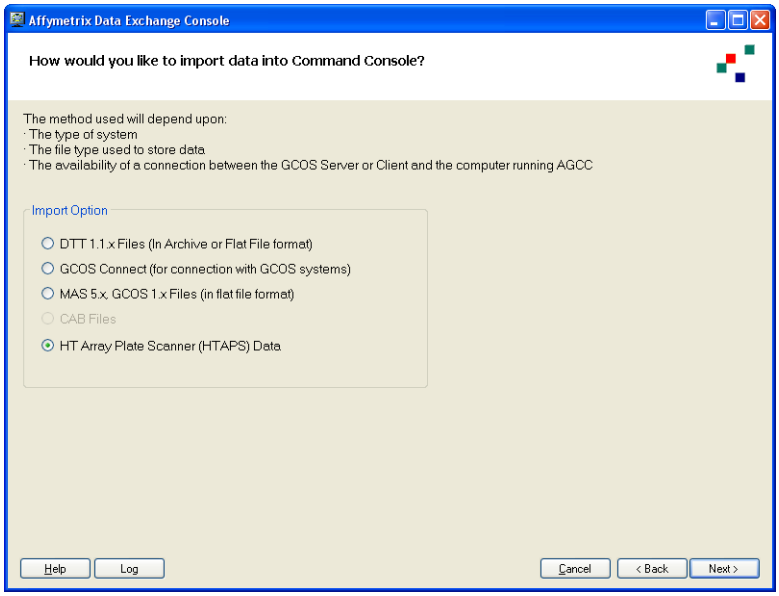
Affymetrix Data Exchange Console
How would you like to import data into Command Console?
The method used will depend upon:
· The type of system
· The file type used to store data
· The availability of a connection between the GCOS Server or Client and the computer running AGCC
Import Option
DTT 1.1.x Files (In Archive or Flat File format)
GCOS Connect (for connection with GCOS systems)
MAS 5.x, GCOS 1.x Files (in flat file format)
CAB Files
HT Array Plate Scanner (HTAPS) Data
Help
Log
Cancel
< Back
Next >

Figure 3.1 Select Import Method

- DTT 1.1.x Files (In Archive or Flat File format)

  Select this option to transfer in data from DTT Archive files or as a set of flat files with associated XML files. The flat files or DTT Archive file must have the associated XML file(s) with the GCOS project and sample information. The DAT/CEL/CHP data will be transferred into AGCC under the appropriate project/sample and the data files will be converted to the appropriate AGCC format.
- GCOS Connect: Direct connection between GCOS and AGCC.

  You can use this to transfer data directly from a GCOS Server into AGCC or from a GCOS Client installation on the same computer as the AGCC installation. The DAT/CEL/CHP data will be transferred into AGCC under the appropriate project/sample and the data files will be converted to the appropriate AGCC format.

> **IMPORTANT:** **If you wish to connect to a GCOS server with DEC, you must install DEC for GCOS Server on the GCOS server. If DEC for GCOS Server is not installed on the server, you will see a permission error when you try to connect. See *Installing DEC for GCOS Server on the GCOS Server* *on page 19* of the *AGCC 2.0 Installation Instructions* for more information.**

- MAS 5.x, GCOS 1.x Data

  Select this option to transfer in DAT/CEL/CHP data that do not have the XML file with the GCOS project and sample information. Other information required for transfer IN will be regenerated using the information available in the DAT/CEL headers. The data is copied to the user provided folder.
- HT Array Plate Scanner (HTAPS) Data

  Select this option to transfer in DAT/CEL/CHP data for an array plate. Other information required for transfer IN will be regenerated using the information available in the DAT/CEL headers. The data is copied to the user provided folder.

To select an import method:

1. Select the format in the list.
2. Click **Next**.

   The appropriate Select Data File window opens.

Importing DTT Files

DTT files are a GCOS interchange format created with the Data Transfer Tool; they use an XML file to transfer the sample and experiment data for the DAT, CEL, and CHP data files.

When exporting data from GCOS in DTT format for import using DEC, you can:

- Place the XML and data files into a zip file for easy transport
- Transfer the separate XML and data files in flat file format.

Selecting Data File

The Select File window (Figure 3.2) allows you to select the DTT or XML files for import.

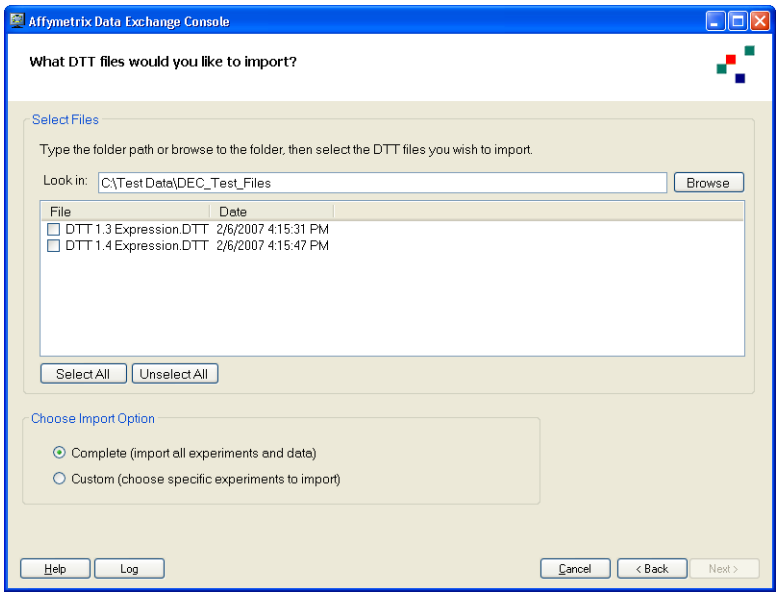

Figure 3.2 Step 2: Select Files

To select DTT or XML files for import:

1. Select the location by typing the path into the Location box, using path completion; or
   - **A.** Click the **Browse** button.
     The Browse for Folder window opens (Figure 3.3).

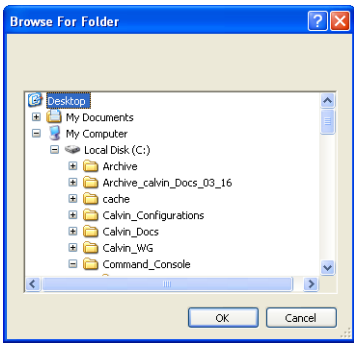

Figure 3.3 Browse for Folder window

   - **B.** Browse to the folder with the files to be imported.
   - **C.** Click **OK**.

A list of the available DTT or XML files is displayed in the box in the Select Files window (Figure 3.4).

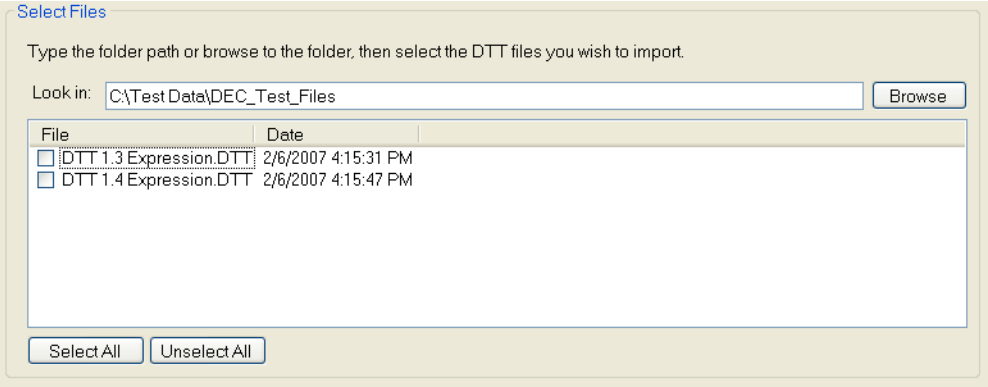

Figure 3.4 DTT or XML files for selection

Each DTT or XML file has a checkbox next to it.

2. Select the checkbox(es) for the file(s) you with to import.
   You can also make and clear selections using the **Select All** and **Unselect All** buttons.
3. Choose Import options:
   - Complete (Import all experiments and data)
     If you select this option, the Select Save Location window (see page 17) opens when you click **Next**.
   - Custom (choose specific experiments to import)
     If you select this option, the Select Experiments window opens when you click **Next**.
4. Click **Next**.
   The appropriate window opens.

Selecting Experiments

The Select Experiments window (Figure 3.5) allow you to select individual experiments in the DTT or XML files.

The list displays the following information:

Experiment	Experiment name.
File Location	DTT file name.
User	User who created the sample in GCOS.
Date	Date of file creation.
Assay Type	General type of array: • Mapping • Expression • Universal • Resequencing
Probe Array Type	Type of probe array analyzed.

You can sort the list order by clicking in the header of the column you wish to sort by.

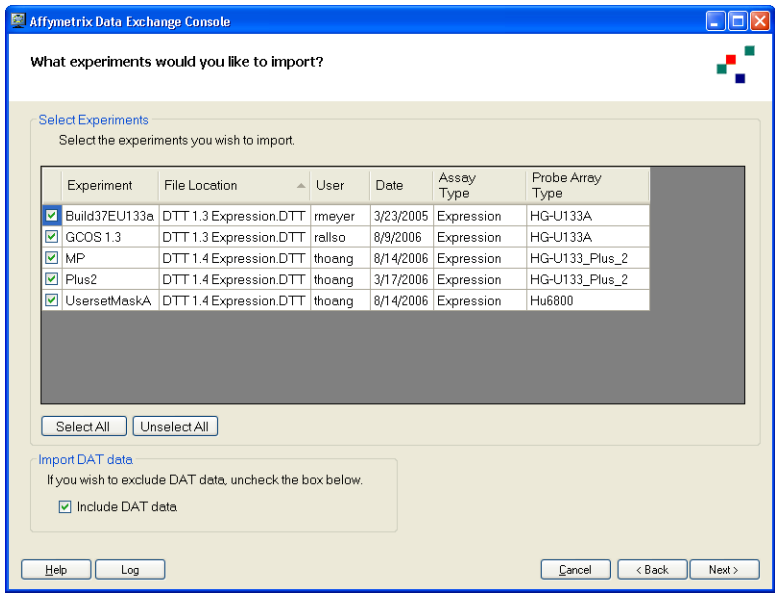

Figure 3.5 DTT Import, Select Experiments

To select experiments:

1. Select the checkbox(es) for the experiment(s) you with to import.
   You can also make and clear selections using the Select All and Unselect All buttons.
2. Deselect **Include DAT data** if you wish to exclude the DAT data from the import.
3. Click **Next**.
   The Select Location window opens (see below).

Selecting a Location

The Select Location window (Figure 3.6) allows you to select a location to place the imported data.

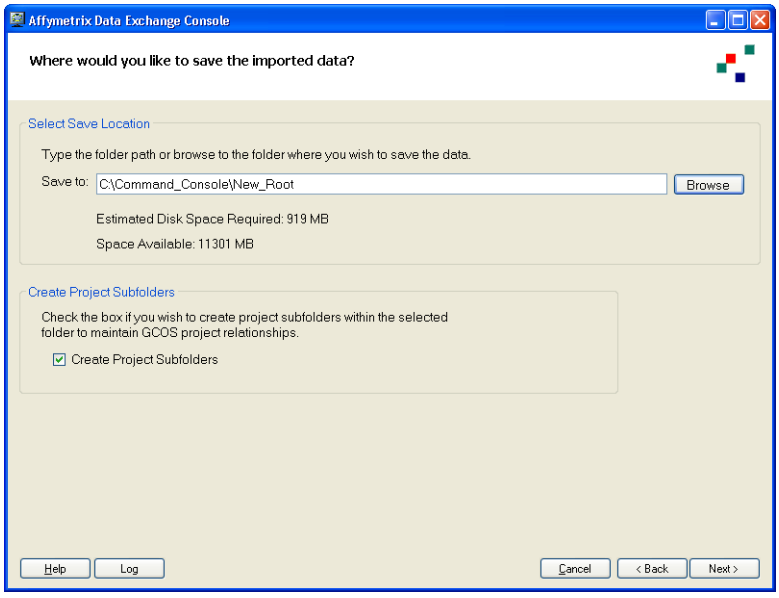
Affymetrix Data Exchange Console
Where would you like to save the imported data?
Select Save Location
Type the folder path or browse to the folder where you wish to save the data.
Save to:
C:\Command_Console\New_Root
Browse
Estimated Disk Space Required: 919 MB
Space Available: 11301 MB
Create Project Subfolders
Check the box if you wish to create project subfolders within the selected folder to maintain GCOS project relationships.
Create Project Subfolders
Help
Log
Cancel
< Back
Next >

Figure 3.6 DTT Import, Select Save Location

1. Select the location by typing the path into the Location box, using path completion; or
   A. Click the **Browse** button.
      The Browse for Folder window opens (Figure 3.7).

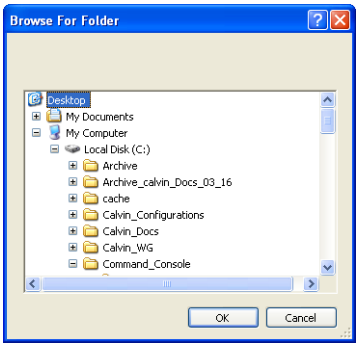

Figure 3.7 Browse for Folder window

   B. Browse to the folder with the files to be imported.
   C. Click **OK**.
2. Select **Create Project Subfolders** if you wish to create project subfolders.

NOTE: This option is not available when importing MAS 5.x or GCOS flat files.

3. Click **Next**.
   The Review Data window opens.

Reviewing Data

The Review Data window (Figure 3.8) displays a list of the files to be imported with the following information:

Status Icon Indicator of the status of the import:
- No conflicts or conflicts resolved.
- Unresolvable conflict (files will be imported on a file by file basis).
- Error condition (no files will be imported).

Sample AGCC Sample file name, based on the GCOS Experiment name.

Type Type of data (CEL, DAT, CHP, GRD, JPG)

Conflict Brief description of the conflict, if any.

Resolution Automatic resolution to the conflict.

Error Description of a condition that cannot be resolved.

Items with errors are discarded on a file by file level (CEL, CHP, etc.).

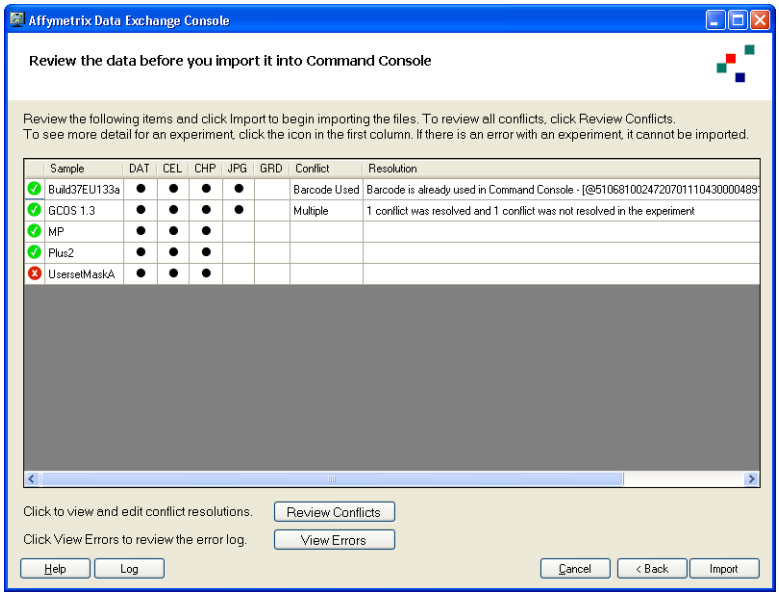
Affymetrix Data Exchange Console
Review the data before you import it into Command Console
Review the following items and click Import to begin importing the files. To review all conflicts, click Review Conflicts.
To see more detail for an experiment, click the icon in the first column. If there is an error with an experiment, it cannot be imported.
Sample DAT CEL CHP JPG GRD Conflict Resolution
Build37EU133a Barcode Used Barcode is already used in Command Console - [@5106810024720701110430000489
GCOS 1.3 Multiple 1 conflict was resolved and 1 conflict was not resolved in the experiment
MP
Plus2
UsersetMaskA
Click to view and edit conflict resolutions.
Review Conflicts
Click View Errors to review the error log.
View Errors
Help
Log
Cancel
< Back
Import

Figure 3.8 DTT Import, Review Data

Conflicts arise because of discrepancies between the data in the different GCOS files being imported and between the data being imported and data already in AGCC.

Click in a row to display the Detailed Validation Results for the experiment being imported (Figure 3.9).

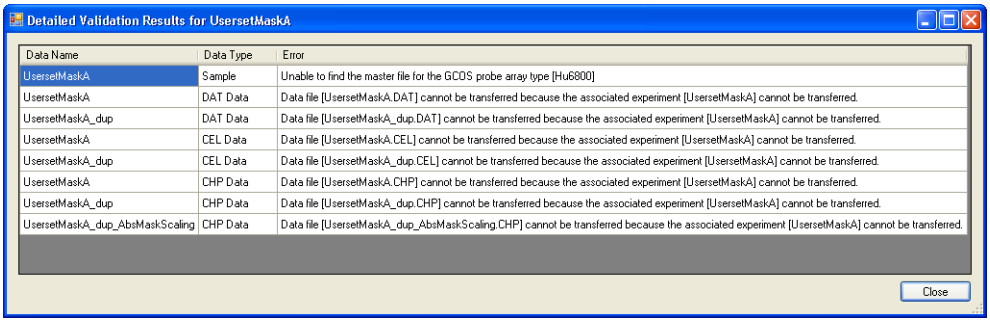
Detailed Validation Results for UsersetMaskA
Data Name Data Type Error
UsersetMaskA Sample Unable to find the master file for the GCOS probe array type [Hu6800]
UsersetMaskA DAT Data Data file [UsersetMaskA.DAT] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA_dup DAT Data Data file [UsersetMaskA_dup.DAT] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA CEL Data Data file [UsersetMaskA.CEL] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA_dup CEL Data Data file [UsersetMaskA_dup.CEL] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA CHP Data Data file [UsersetMaskA.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA_dup CHP Data Data file [UsersetMaskA_dup.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA_dup_AbsMaskScaling CHP Data Data file [UsersetMaskA_dup_AbsMaskScaling.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
Close

Figure 3.9 Detailed Validation Results

The list displays the following information about the experiment:

Data Name Name of file.

Data Type Type of file being imported or created.

Error Brief description of the condition that cannot be resolved, if any.

To review the data for import:

1. Review the list for conflicts.

   Conflicts will automatically be resolved, but you can override the automatic resolution using the Review Conflicts window.

2. Click the **Review Conflicts** button.

   The Review Conflicts window opens (Figure 3.10).

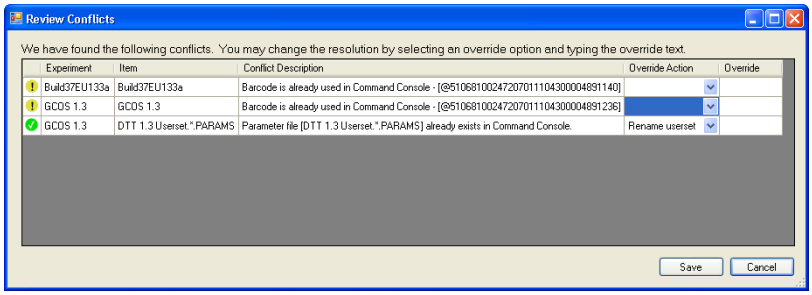

Figure 3.10 Review Conflicts

The window displays a list of conflicts with the following information:

Status Icon Indicator of the status of the import:

- No conflicts or conflicts resolved.
- Unresolvable conflict (files will be imported on a file by file basis).
- Error condition (no files will be imported).

Experiment Name of imported experiment.

Items Particular file causing problem (DAT, template, etc.).

Conflict Description Description of the problem

Override Action Resolution automatically selected.

Override Enter a new name for the file to override the automatically created name.

3. Enter new names in the Override column if desired.
4. Click **Save** to close the Review Conflicts window and enter the new file names.
5. Click **View Errors** to review files that cannot be imported at all.
   The View Errors window opens (Figure 3.11).

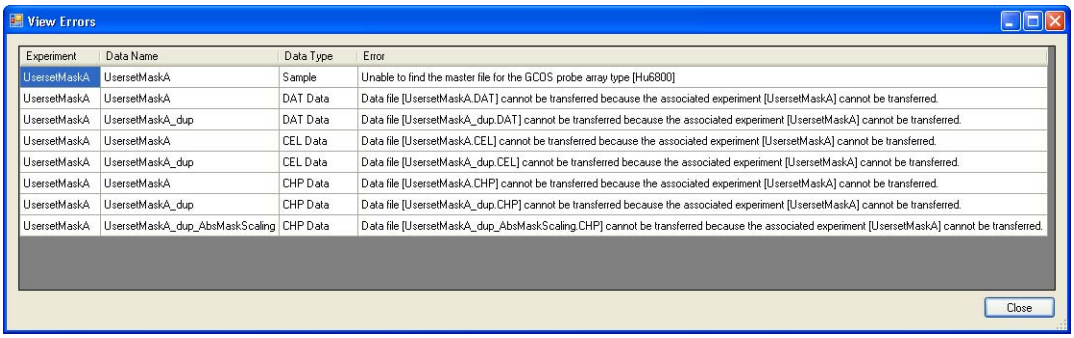
View Errors
Experiment
Data Name
Data Type
Error
UsersetMaskA
UsersetMaskA
Sample
Unable to find the master file for the GCOS probe array type [Hu6800]
UsersetMaskA
UsersetMaskA
DAT Data
Data file [UsersetMaskA.DAT] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA_dup
DAT Data
Data file [UsersetMaskA_dup.DAT] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA
CEL Data
Data file [UsersetMaskA.CEL] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA_dup
CEL Data
Data file [UsersetMaskA_dup.CEL] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA
CHP Data
Data file [UsersetMaskA.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA_dup
CHP Data
Data file [UsersetMaskA_dup.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
UsersetMaskA
UsersetMaskA_dup_AbsMaskScaling
CHP Data
Data file [UsersetMaskA_dup_AbsMaskScaling.CHP] cannot be transferred because the associated experiment [UsersetMaskA] cannot be transferred.
Close

Figure 3.11 View Errors

The window displays a list of files and data with the following information:

Experiment Name of imported experiment.

Data Name Particular file causing problem (DAT, template, etc.).

Data Type Type of file.

Error Reason the file cannot be imported.

6. Click **Close** in the View Error window.
7. Click **Import**.
   The Import Status window opens (Figure 3.12).

Tracking Import Status

The Import Status window (Figure 3.12) displays a list of the imported files with information about whether the import was successful or not.

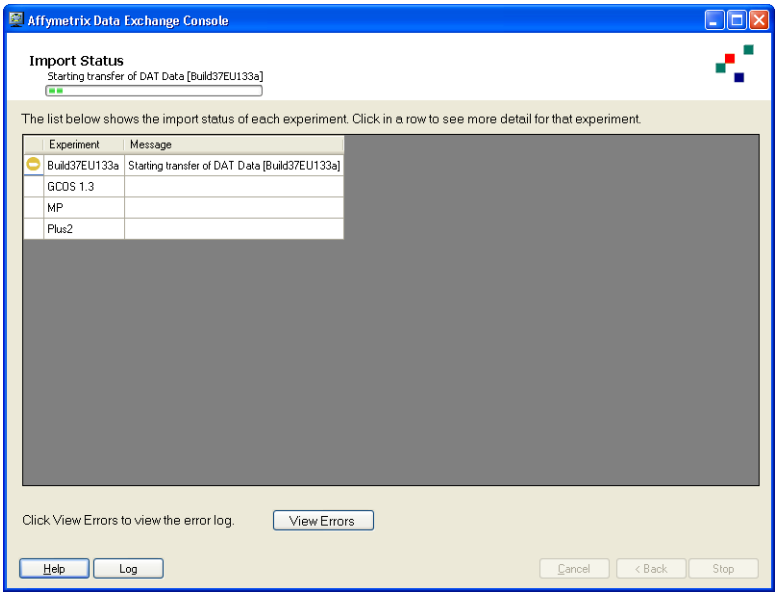

Figure 3.12 Import Status window

The screen displays the following columns:

Status Icon Indicator of the status of the import:

- No conflicts or conflicts resolved.
- In progress.
- Unresolvable conflict, error condition (File will not be imported).

Experiment Name of imported experiment

Message Particular file causing problem (DAT, template, etc.)

Click in a row to display the Detailed Migration Results window for the experiment being imported (Figure 3.13).

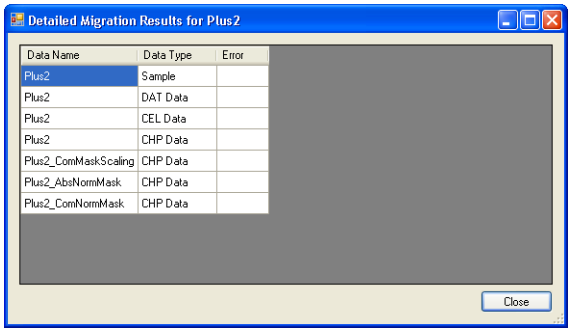

Figure 3.13 Detailed Migration Results window

The window lists the following information about the experiment:

Data Name Name of file.

Data Type Type of file being imported or created.

Error Brief description of the condition that cannot be resolved, if any.

After the import is finished:

- Click **View Errors** to view the error log; or
  Click **Start Again** to import more files; or
  Click **Finish** to close the DEC window.

The imported data is visible in the AGCC Data folder.

Importing From GCOS Connect

GCOS Connect allows you to transfer data from a GCOS system on another computer or on the computer with AGCC in one step. You will not have to export the files from GCOS using DTT.

NOTE: DTT must be installed on the GCOS system you are importing files from to use GCOS Connect.

Selecting a GCOS System

The Select GCOS System window (Figure 3.14) allows you to select a GCOS system from which to import data.

IMPORTANT: If you wish to connect to a GCOS server with DEC, you must install DEC for GCOS Server on the GCOS server. If DEC for GCOS Server is not installed on the server, you will see a permission error when you try to connect. See *Installing DEC for GCOS Server on the GCOS Server* on page 19 of the *AGCC 1.1 Installation Instructions* for more information.

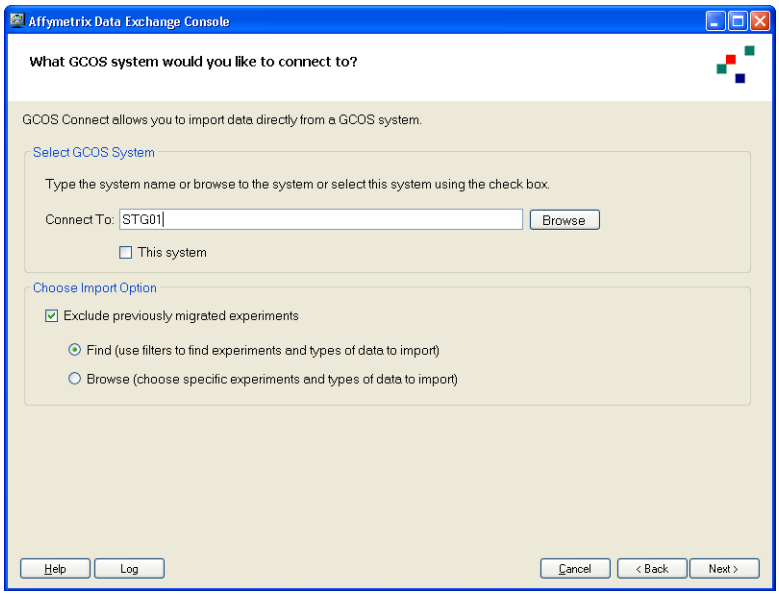
Affymetrix Data Exchange Console
What GCOS system would you like to connect to?
GCOS Connect allows you to import data directly from a GCOS system.
Select GCOS System
Type the system name or browse to the system or select this system using the check box.
Connect To: STG01
Browse
This system
Choose Import Option
Exclude previously migrated experiments
Find (use filters to find experiments and types of data to import)
Browse (choose specific experiments and types of data to import)
Help
Log
Cancel
< Back
Next >

Figure 3.14 Select GCOS System

To select a GCOS system:

1. Enter a server name in the Connect To box; or
   Browse to the location of the server.

   A. Click the **Browse** button.

   The Browse for Folder window opens (Figure 3.15).

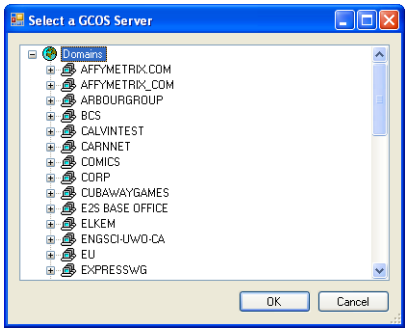

Figure 3.15 Browse for Folder window

   B. Browse to the folder with the files to be imported

   C. Click **OK**.

   The server name appears in the Connect To box. (Figure 3.14).

2. You can select an import option to limit the files on the GCOS server that are displayed in DEC:
   - Exclude previously migrated files
   - Find (use filters to find experiments and types of data to import)
     If you select this option, the Filter window appears when you click **Next**.
   - Browse (display all experiments on the GCOS server)
   - If you choose this option, the Select Experiments window appears when you click **Next.**

3. Click **Next**.
   The appropriate window opens.

Filtering the Experiments

You can use this window to filter experiments displayed in the select experiments window (Figure 3.16).

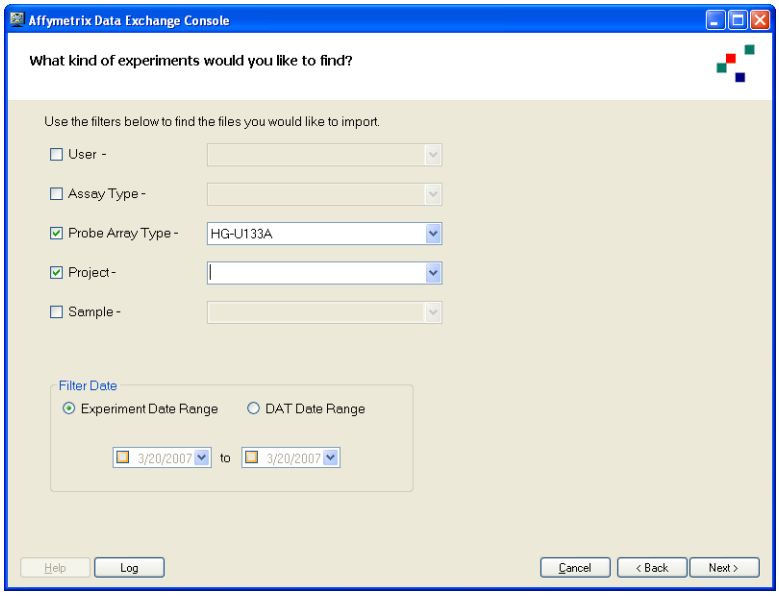

Figure 3.16 Filter Data

To filter experiments:

1. Select the GCOS attribute types and attributes from the following items:

User	User who created the sample in GCOS
Assay Type	General type of array: • Mapping • Expression • Universal • Resequencing
Probe Array Type	Type of probe array analyzed.
Project	Project that the probe array was listed under.
Sample	Sample Name used for the probe array.

   A. Click the checkbox next to the attribute name.

   B. Select an attribute from the drop down list.

2. Filter by dates by doing the following:

   A. Select an option from:

Experiment Date Range	Dates when the experiment was created.
DAT Date Range	Dates when the array was scanned.

B. Select a date or range of dates for file creation:

1) Select the **From** checkbox.

2) Click the arrow at the date (displays the current date).

A calendar for the current month appears (Figure 3.17).

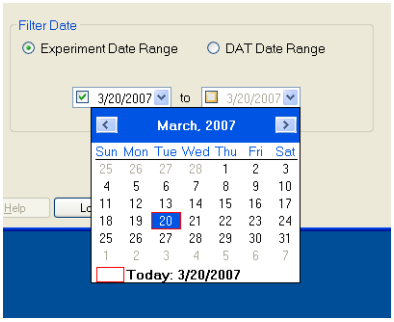

Figure 3.17 From Calendar

3) Select a date for the start of the range. You can move from month to month by clicking the < and > buttons.

If you only select a start date, the filter will display only the files created on or after that date.

To select a range of dates:

- Select a date for the end of the range.

  If you only select an end date, the filter will display only the files created on or before that date.

3. Click **Next**.

The Select Experiments window appears.

Selecting the Experiments

The Select Experiments window (Figure 3.18) allow you to select individual experiments in the GCOS system.

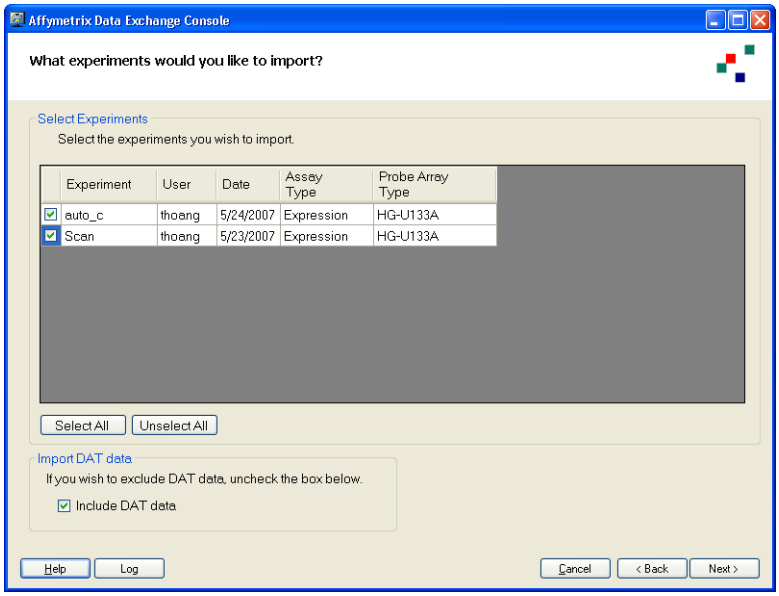

Figure 3.18 GCOS Connect Import, Select Experiments

The list displays the following information:

Experiment	Experiment name.
User	User who created the sample in GCOS.
Date	Date of file creation.
Assay Type	General type of array: • Mapping • Expression • Universal • Resequencing
Probe Array Type	Type of probe array analyzed.

You can sort the list order by clicking in the header of the column you wish to sort by.

To select experiments:

1. Select the checkbox(es) for the experiment(s) you with to import.
   You can also make and clear selections using the Select All and Unselect All buttons.

NOTE: Files to which you do not have permission to access on GCOS are displayed in this list. An error message will be displayed when you go to the Review page.

   You can also make and clear selections using the Select All and Unselect All buttons.
2. Deselect **Include DAT data** if you wish to exclude the DAT data from the import.
3. Click **Next**.
   The Select Location window opens.

Selecting a Location

See *Selecting a Location* on page 17.

Reviewing Data

If you are trying to import files that you do not have access to in the GCOS Server, you will see the following error message (Figure 3.19):

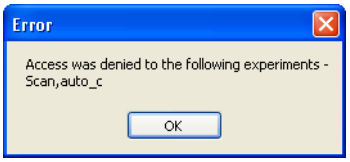

Figure 3.19 Access denied message

You will need to contact the administrator for the GCOS Server to get access to the files.

See *Reviewing Data* on page 18 for information about checking for data conflicts and errors.

Importing Data

See *Tracking Import Status* on page 21.

Importing MAS/GCOS Flat Files

You may wish to import MAS and GCOS flat files from earlier versions of GCOS. These files will not have the sample information that is available with DTT files or GCOS Connect.

When you are importing the DAT, CEL, and CHP files by themselves, DEC will create a Sample (.ARR) file for each group of data files. DEC will group data files using file names and embedded parent file information, and you can change the grouping if necessary to accommodate special circumstances.

Selecting Data File

The Select File window (Figure 3.20) allows you to select the data files for import.

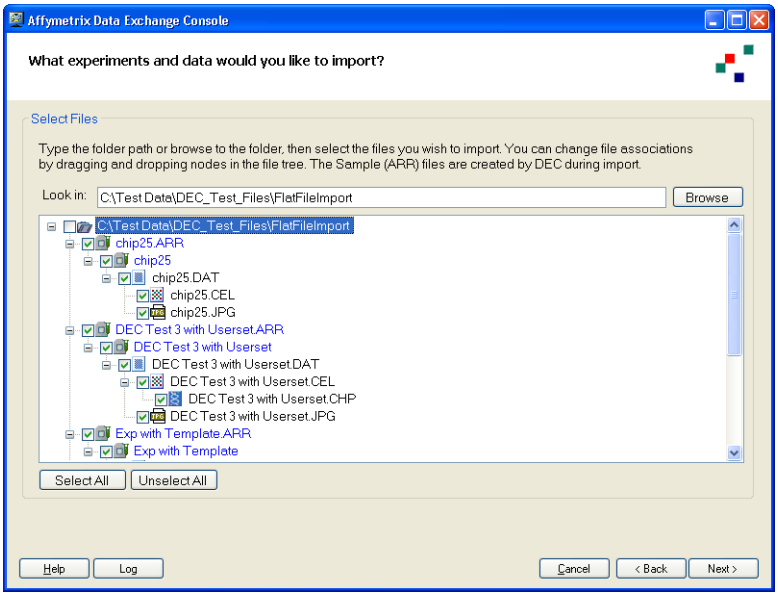
Affymetrix Data Exchange Console
What experiments and data would you like to import?
Select Files
Type the folder path or browse to the folder, then select the files you wish to import. You can change file associations by dragging and dropping nodes in the file tree. The Sample (ARR) files are created by DEC during import.
Look in:
C:\Test Data\DEC_Test_Files\FlatFileImport
Browse
C:\Test Data\DEC_Test_Files\FlatFileImport
chip25.ARR
chip25
chip25.DAT
chip25.CEL
chip25.JPG
DEC Test 3 with Userset.ARR
DEC Test 3 with Userset
DEC Test 3 with Userset.DAT
DEC Test 3 with Userset.CEL
DEC Test 3 with Userset.CHP
DEC Test 3 with Userset.JPG
Exp with Template.ARR
Exp with Template
Select All
Unselect All
Help
Log
Cancel
< Back
Next >

Figure 3.20 Select File

The page displays a list of the available files in a hierarchical tree. The Data files are arranged in groups, based on the file names and the embedded parent file information. It also displays the Sample file and Array Name(s) that will be created for the files in AGCC.

To select files for import:

1. Select the location by typing the path into the Location box, using path completion; or
   A. Click the **Browse** button.
      The Browse for Folder window opens (Figure 3.21).

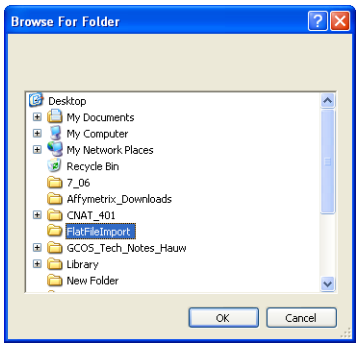

Figure 3.21 Browse for Folder window

B. Browse to the folder with the files to be imported.

C. Click **OK**.

A list of the available flat files is displayed in the box in the Select Files window (Figure 3.22).

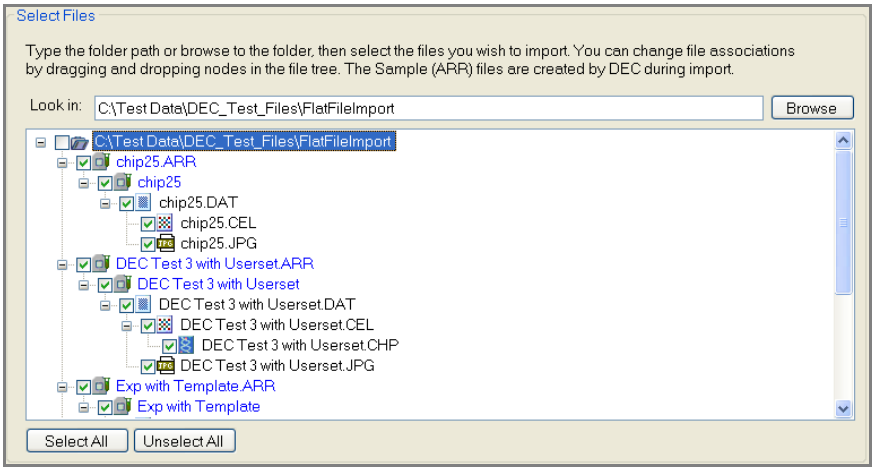
Select Files
Type the folder path or browse to the folder, then select the files you wish to import. You can change file associations by dragging and dropping nodes in the file tree. The Sample (ARR) files are created by DEC during import.
Look in:
C:\Test Data\DEC_Test_Files\FlatFileImport
Browse
C:\Test Data\DEC_Test_Files\FlatFileImport
chip25.ARR
chip25
chip25.DAT
chip25.CEL
chip25.JPG
DEC Test 3 with Userset.ARR
DEC Test 3 with Userset
DEC Test 3 with Userset.DAT
DEC Test 3 with Userset.CEL
DEC Test 3 with Userset.CHP
DEC Test 3 with Userset.JPG
Exp with Template.ARR
Exp with Template
Select All
Unselect All

Figure 3.22 files for selection

Each flat file has a checkbox next to it.

The data files (DAT, CEL, and CHP) are grouped by file name and embedded parent file information and associated with a Sample (.ARR) file that will be created upon import. You can select Sample and data files for import, and move data files using drag and drop from one group to another to accommodate special circumstances.

If you are importing the following flat files (Figure 3.23):

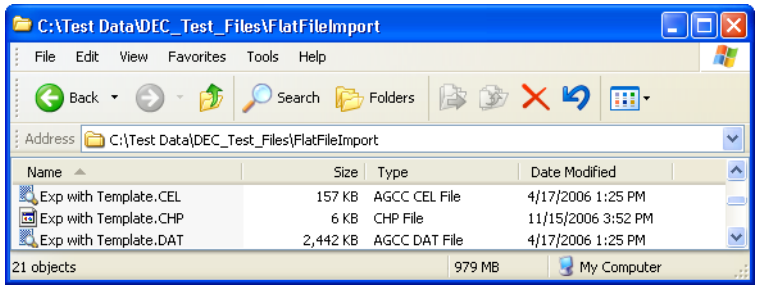

Figure 3.23 Data files to be imported

You will see the following file group in the Select Files list (Figure 3.24):

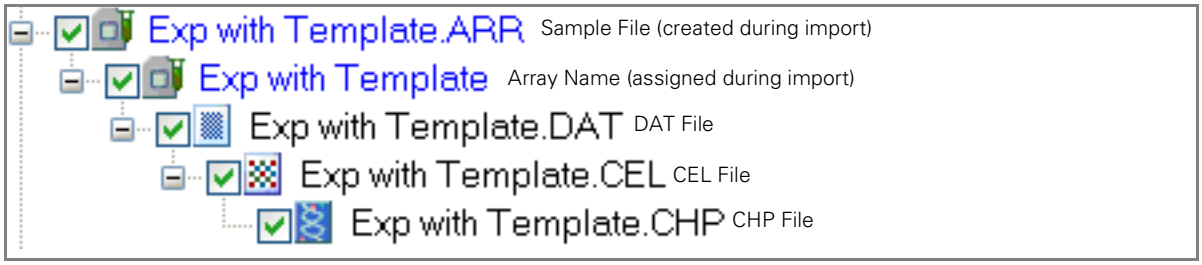

Figure 3.24 Sample file with Array name and data files

2. Select the checkbox(es) for the file(s) you with to import. Selecting the checkbox for a parent file automatically selects all of the child file checkboxes; you can deselect individual child checkboxes if desired.

   You can also make and clear selections using the **Select All** and **Unselect All** buttons.

3. Move arrays and data files from one sample file to another if necessary to get the relationships between files correct.

   For example, if you have run duplicate arrays for a sample in GCOS (Figure 3.25) the duplicate arrays will be initially set up with separate array names and Sample files.

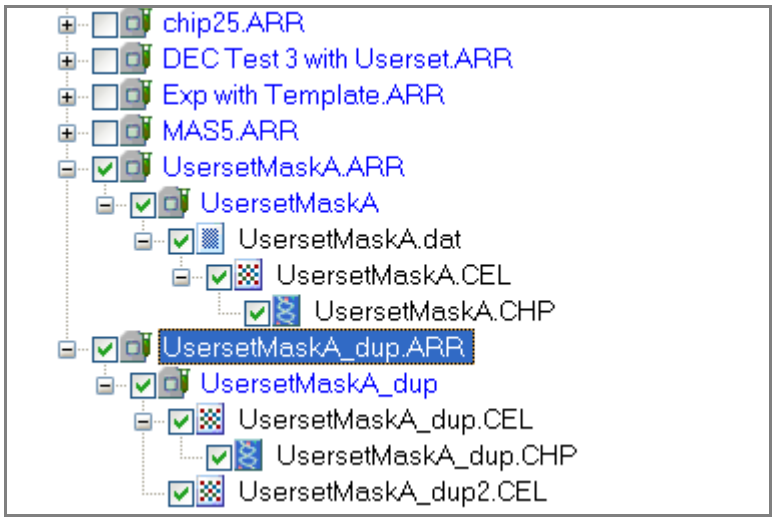

Figure 3.25 Duplicate arrays

You can consolidate the arrays under one Sample file by moving the duplicate array name (and all its data files) to the first Sample file (Figure 3.26).

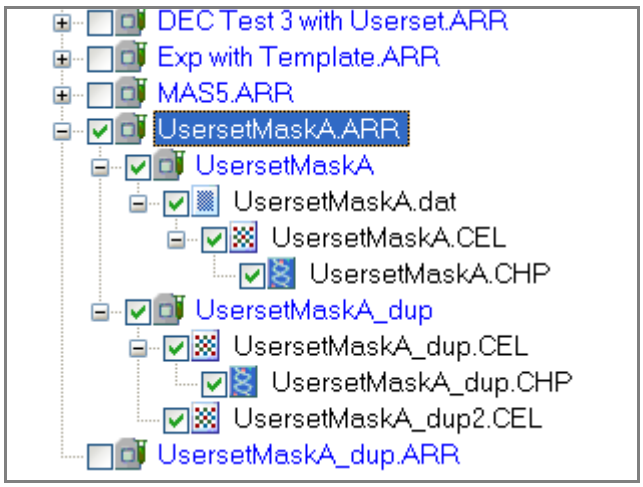

Figure 3.26 UsersetMaskA_dup moved to UsersetMaskA.ARR sample file

4. Click **Next**.

   The Select Location window opens.

Selecting a Location

See *Selecting a Location* on page 17.

Reviewing Data

See *Reviewing Data* on page 18.

Tracking Import Status

See *Tracking Import Status* on page 21

Importing HT Array Plate Scanner Data

HT Array Plate Scanner (HTAPS) data can be imported into AGCC 2.0 for management and use. Experiment information is not imported.

HT Array Plate Scanner

The HT Array Plate Scanner window (Figure 3.27) allows you to select a set of data for one or more array plates to import.

NOTE: We recommend moving all the peg level data (miniDAT, CEL and CHP) files together. Reconnector does not support re-establishing relationships between HT data files if they are imported separately into AGCC.

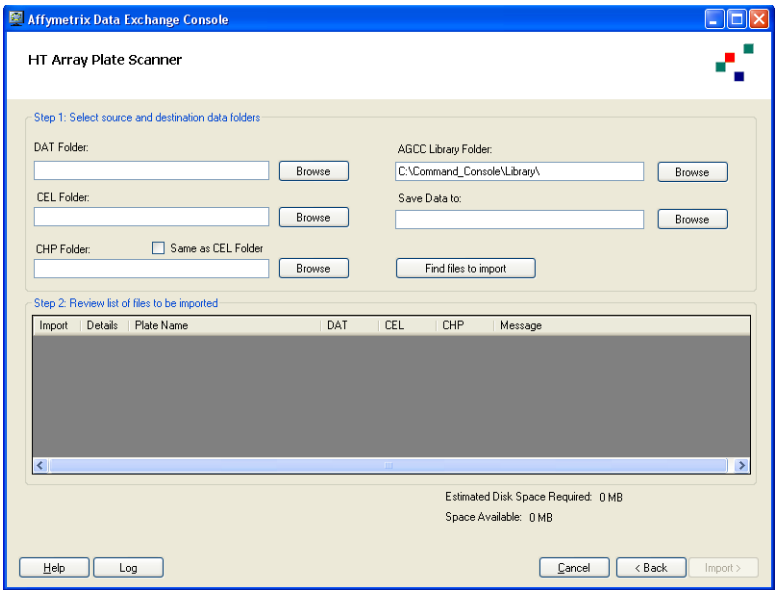

Figure 3.27 HT Array Plate Scanner window

To import GCOS HT array plate data:

1. Select the location of:
   - DAT files
   - CEL files
   - CHP Files

   Type the path into the appropriate box, using path completion; or

   A. Click the **Browse** button.

      The Browse for Folder window opens (Figure 3.21).

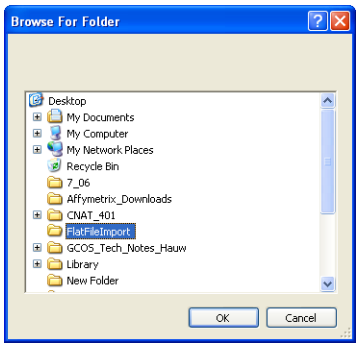

Figure 3.28 Browse for Folder window

B. Browse to the folder with the files to be imported.

C. Click **OK**.

2. Enter the location of the AGCC Library Folder by typing the path into the Location box or using the **Browse** Button.
3. Enter the destination location for the imported files in the Save Data To box by typing the path into the box or using the **Browse** button.
4. Click **Find Files to Import**.

   The window displays a list of files to be imported in the **Step 2: Review list of files to be imported** list.

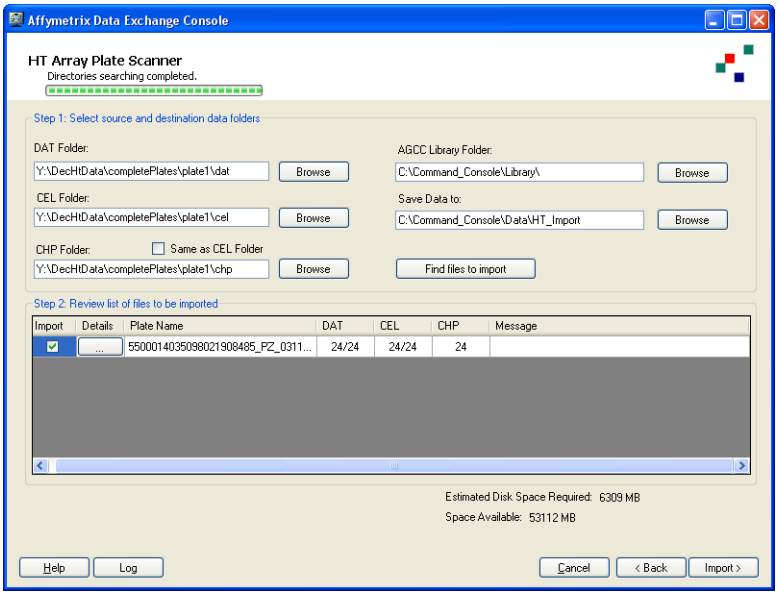

Figure 3.29 Review List displayed

The Review list displays (Figure 3.30):

Import Select the checkbox to import the files for the array plate.

Details Click the Details button to display the Detailed Plate Results for dialog box (Figure 3.33).

Plate Name Unique Plate identifier as specified in GCOS.

DAT Number of DAT files to be imported/standard number of DAT files for this array plate type.

CEL Number of CEL files to be imported/standard number of CEL files for this array plate type.

CHP Number of CHP files to be imported (can have multiple CHP files for a CEL file).

Message Additional information, notice of problems, etc.

Problems with the files are displayed in the list:

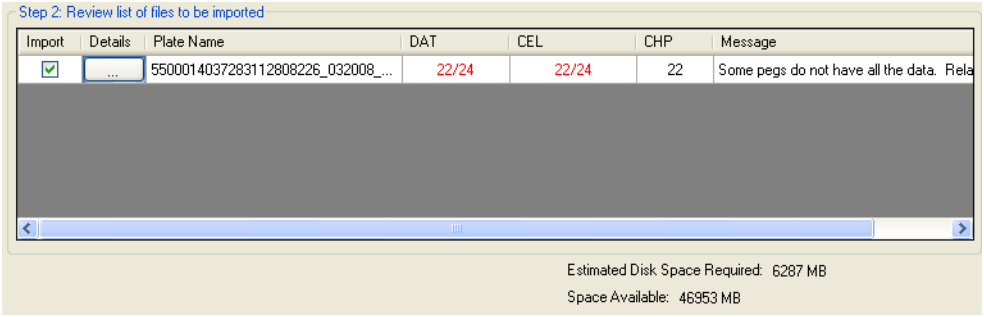

Figure 3.30 Review List with Missing DAT and CEL files

An estimate of the disk space required for import and the disk space availabe are displayed below the list.

If there are problems, the file types that are incorrect are displayed in red and error messages are displayed.

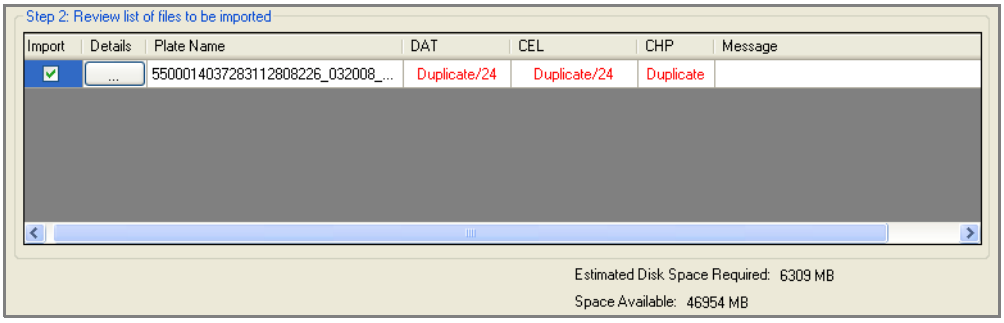

Figure 3.31 Error message for duplicate files

If these error messages appear, you will need to go to the folder with the files you wish to import and resolve the problem there before importing. The resolution may involve:

- Finding missing files
- Moving duplicate files

You can view a detailed list of the files for individual arrays in the Detail Plate Results list.

- Click the Details button in the Review List (Figure 3.32)

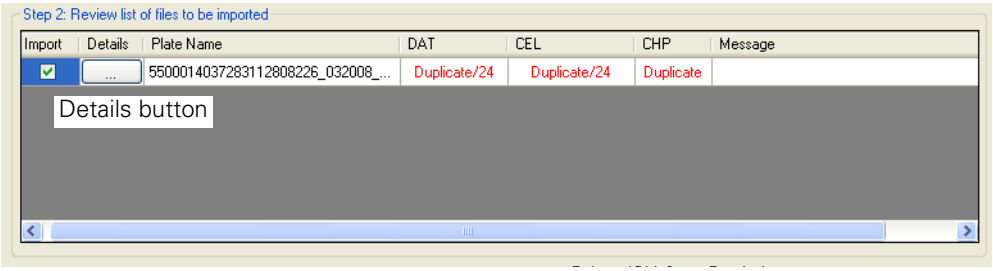

Figure 3.32 Detailed button in the Results List

The Detailed Plate Results dialog box (Figure 3.33) opens.

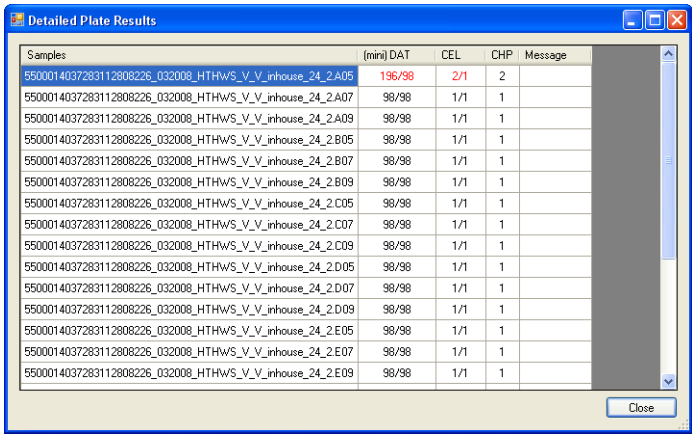

Figure 3.33 Detailed Plate Results dialog box for duplicate peg info

The dialog box displays:

Samples Name for Sample (ARR) file to be created.

(mini) DAT Number of mini-DAT files to be imported for the array/number of mini-DAT files that are standard for the array plate type.

CEL Number of CEL files to be imported.

CHP Number of CHP files

Message Additional information, notice of problems, etc.

5. Review the data to be imported and click the **Import** button.
The import begins and the Import Progress window opens (Figure 3.34).

Import Progress

The Import Status window (Figure 3.34) displays a list of the imported files with information about whether the import was successful or not.

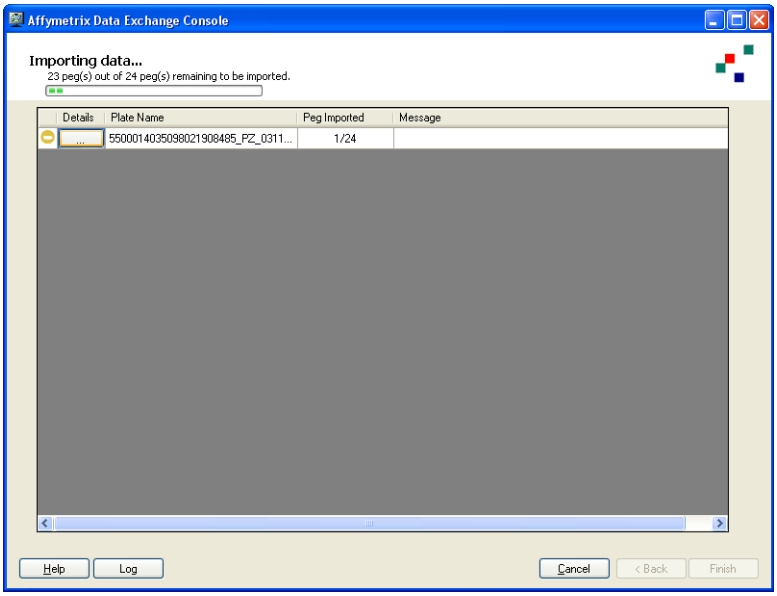

Figure 3.34 Import Progress

The screen displays the following columns:

Status Icon Indicator of the status of the import:

- No conflicts or conflicts resolved.
- In progress.
- Unresolvable conflict, error condition (File will not be imported).

Plate Name Name of the plate.

Peg Imported Number of array peg with data imported/number of total array pegs on array.

Message Particular file causing problem (DAT, template, etc.)

Click on the Details button to display the Detailed Plate Results window (Figure 3.35) for the experiment being imported.

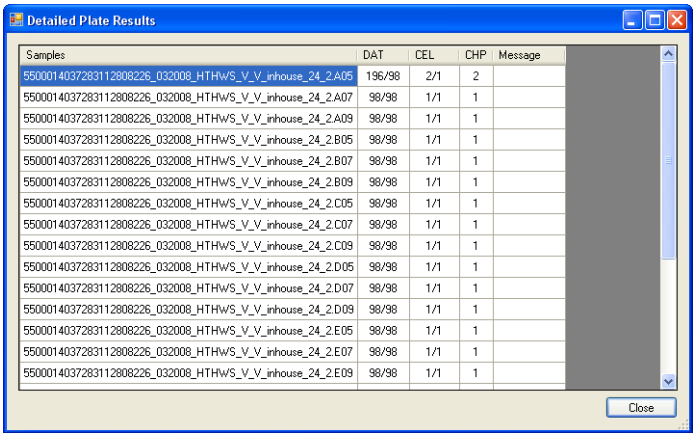

Figure 3.35 Detailed Plate Results window

After the import is finished:

- Click **Back** to import more data; or
  Click **Finish** to close the DEC window.

The imported data is visible in the Data folder.

Chapter 4

DEC EXPORT

The DEC Export option allows you to export data in DTT format for use in a GCOS system.

DEC Export involves the following sets of steps:

1. *Choosing How to Select Files for Export*
2. *Using the Filter* on page 38
3. *Selecting Sample Files for Export* on page 40
4. *Selecting a Location* on page 41
5. *Reviewing Data and Fixing Conflicts* on page 42
6. *Export Status* on page 45

NOTE: **CHP files created with the new expression algorithms (e.g., PLIER, RMA, iterPLIER) and genotyping algorithms (e.g., BRLMM_P, Birdseed) cannot be loaded into GCOS and will not be included in the DTT file.**

NOTE: **HT data from AGCC cannot be imported back into GCOS.**

Choosing How to Select Files for Export

You can select files to export:

- From a filtered set of Sample files.
- From a list of all the Sample files in the AGCC system.

The Export Options window (Figure 4.1) allows you to choose whether to filter the list of files or not.

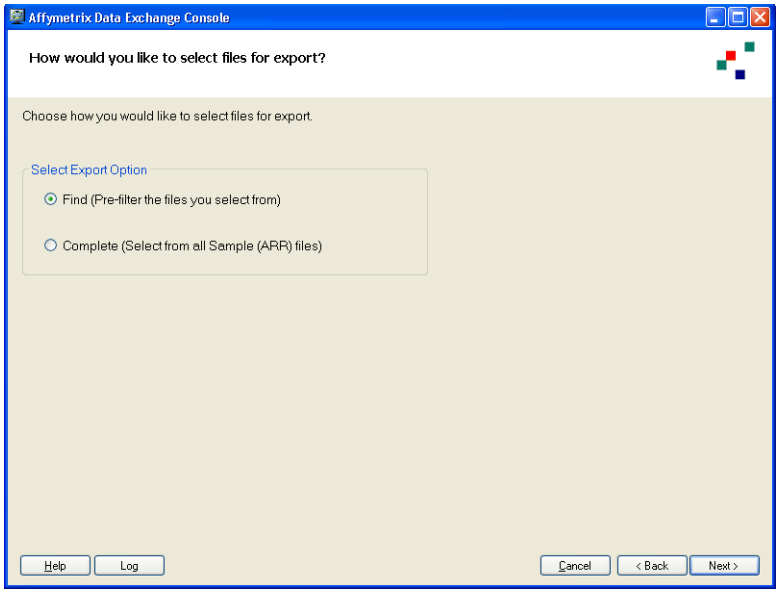
Affymetrix Data Exchange Console
How would you like to select files for export?
Choose how you would like to select files for export.
Select Export Option
Find (Pre-filter the files you select from)
Complete (Select from all Sample (ARR) files)
Help
Log
Cancel
< Back
Next >

Figure 4.1 Select Export Options

To select an option:

1. Select:
   - **Find** to filter the Sample files
     If you select this option, the Filter window opens.
   - **Complete** to select from all sample files and data
     If you select this option, the Select Sample Files window appears.
2. Click **Next**.
   The appropriate window opens.

Using the Filter

The Filter window (Figure 4.2) allows you to filter the Sample files that will be displayed in the Select Array files window.

You can filter by:

- Project
- Dates that Sample (ARR) file or DAT files were created

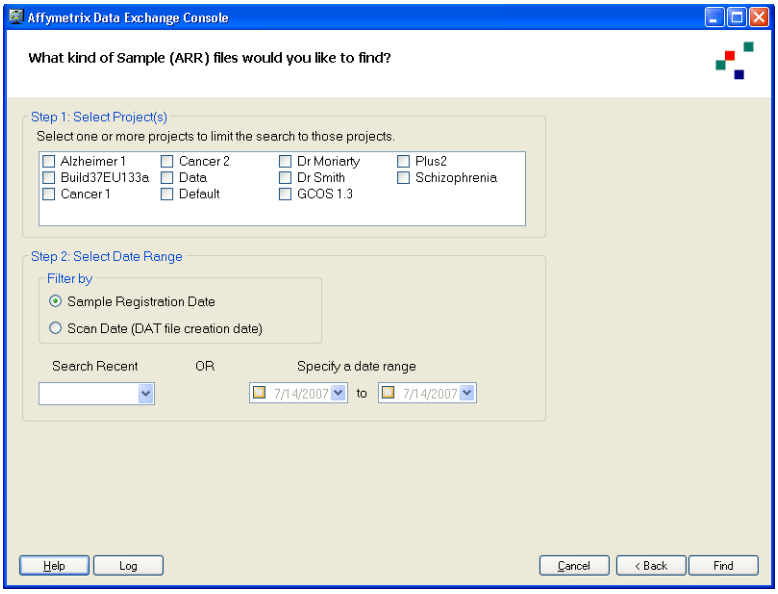

Figure 4.2 Data Filters

To filter:

1. Select a project from the Select Projects list.
2. Select a Date Filter option:

Sample Registration Date	Date the Sample (ARR) file was created.
Scan Date	Date the DAT file was created.

3. Select a date or range of dates for file creation:
   - Select a date limit from the Search Recent drop-down box; or

   A. Click on the left-hand date range box.

   The checkbox is selected and a calendar for the current month appears (Figure 4.3).

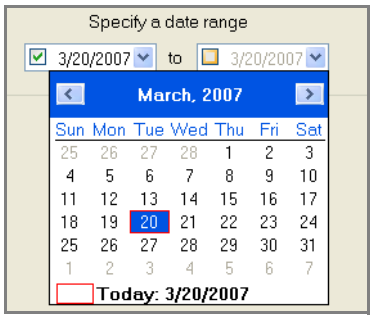

Figure 4.3 From Calendar

   B. Select a date for the start of the range. You can move from month to month by clicking the < and > buttons.

   If you only select a start date, the filter will display only the files created on or after that date.

   To select a range of dates:

- Select a date for the end of the range.

  If you only select an end date, the filter will display only the files created on or before that date.

4. Click **Find**.

   The Select Sample Files window appears.

Selecting Sample Files for Export

The Select Files window (Figure 4.4) displays a list of Sample (ARR) files that can be exported.

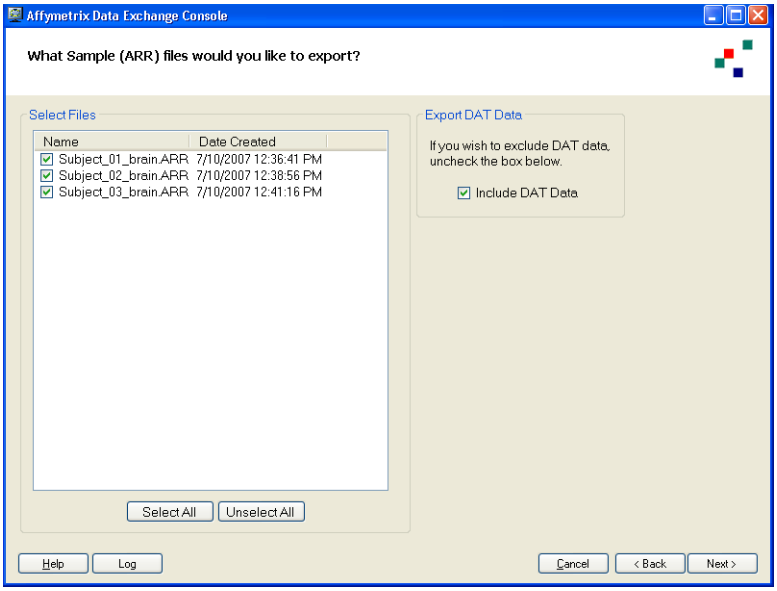

Figure 4.4 Select Data

To select Sample (ARR) files:

1. Select the files for export from the list.

   You can also make and clear selections using the **Select All** and **Unselect All** buttons.

2. Deselect **Include DAT data** if you wish to exclude the DAT data from the export.

3. Click **Next**.

   The Select Location window opens.

Selecting a Location

The Select Location window (Figure 4.5) allows you to select the folder where the DTT file will be placed.

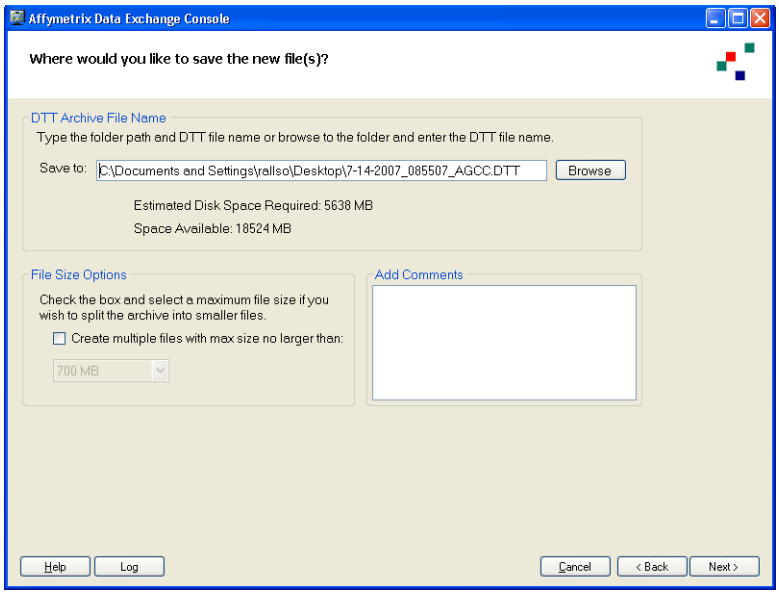
Affymetrix Data Exchange Console
Where would you like to save the new file(s)?
DTT Archive File Name
Type the folder path and DTT file name or browse to the folder and enter the DTT file name.
Save to:
C:\Documents and Settings\rallso\Desktop\7-14-2007_085507_AGCC.DTT
Browse
Estimated Disk Space Required: 5638 MB
Space Available: 18524 MB
File Size Options
Check the box and select a maximum file size if you wish to split the archive into smaller files.
Create multiple files with max size no larger than:
700 MB
Add Comments
Help
Log
Cancel
< Back
Next >

Figure 4.5 Select Location

To select a location for the exported files:

1. Select the location and file name by typing the path and file name into the Location box, using path completion; or

   A. Click the **Browse** button.

   The Save As window opens (Figure 4.6).

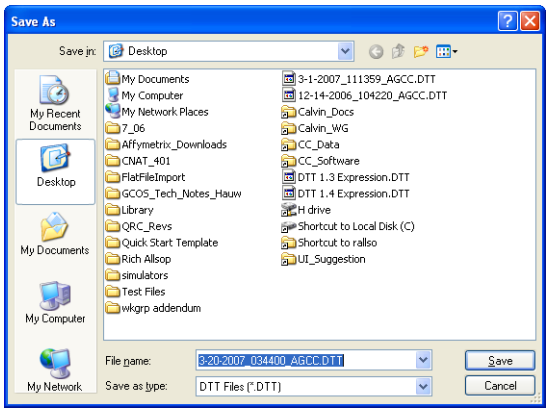

Figure 4.6 Browse (select project file)

   B. Browse through the directory to select a location for the DTT file.

   Enter a new name for the DTT file in the File Name box or use the default name.

   C. In the Save As window, click **OK**.

   Space required and available is displayed.

2. Select **Create Multiple Files** if you wish to split the data into smaller files
3. Select a file size from the drop-down box.

4. Add comments in the Comments box if desired.
5. Click **Next**.
   The Review Data window opens.

Reviewing Data and Fixing Conflicts

The Review Data window (Figure 4.7) displays a list of the files to be exported with the following information:

Status Icon Indicator of the status of the export:
- No conflicts or conflicts resolved.
- Unresolvable conflict (files will be exported on a file by file basis).
- Error condition (File will not be exported).

Experiment Experiment name

Type Type of data (CEL, DAT, CHP, GRD, JPG)

Conflict Brief description of the conflict, if any.

Resolution Automatic resolution to the conflict.

Error Description of a condition that cannot be resolved.

NOTE: CHP files created with the new expression algorithms (e.g., PLIER, RMA, iterPLIER) and genotyping algorithms (e.g., BRLMM_P, Birdseed) cannot be loaded into GCOS and will not be included in the DTT file.

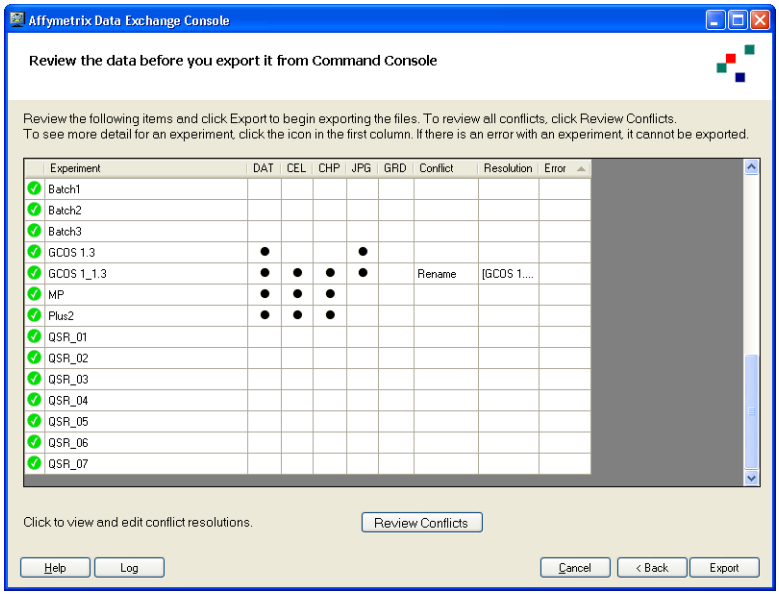

Figure 4.7 DTT export, Review Data

Conflicts arise because of discrepancies between the data in the different AGCC files being exported. Click in a row to display the Detailed Validation Results for the experiment being exported (Figure 4.8).

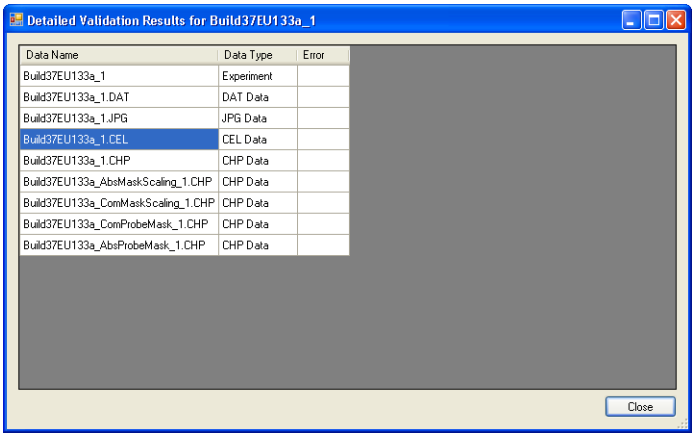

Figure 4.8 Detailed Validation Results

The list displays the following information about the experiment:

Data Name Name of file.

Data Type Type of file being exported or created.

Error Brief description of the condition that cannot be resolved, if any.

To review the data for export:

1. Review the list for conflicts.

   Conflicts will automatically be resolved, but you can override the automatic resolution using the Review Conflicts window.

   CHP files created with the new expression algorithms (e.g., PLIER, RMA, iterPLIER) and genotyping algorithms (e.g., BRLMM_P, Birdseed) cannot be loaded into GCOS and will not be included in the DTT file.

   If this happens, the following notice appears: (Figure 4.9)

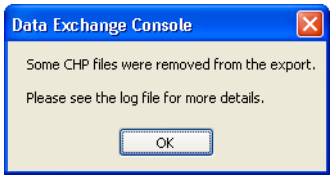

Figure 4.9 CHP file removal notice

2. Click **OK** in the notice.
3. Click the **Review Conflicts** button to override the automatic resolutions.

   The Review Conflicts window opens (Figure 4.10).

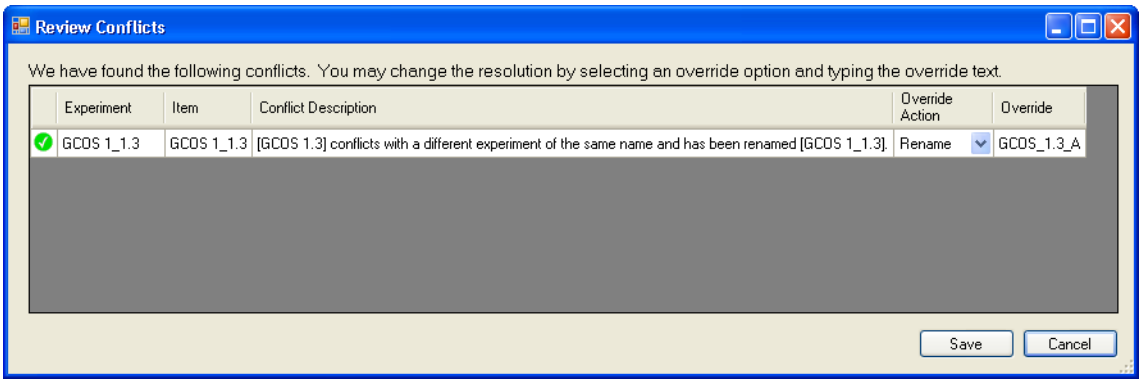

Figure 4.10 Review Conflicts

The window displays a list of conflicts with the following information:

Status Icon Indicator of the status of the export:

No conflicts or conflicts resolved.

Unresolvable conflict (files will be exported on a file by file basis).

Unresolvable conflict, error condition (File will not be exported).

Experiment Name of exported experiment

Items Particular file causing problem (DAT, template, etc.)

Conflict Description Description of the problem

Override Action Resolution automatically selected.

Override Enter a new name for the file to override the automatically created name.

4. Enter new file names as desired.
5. Click **Save** to close the Review Conflicts window and enter the new file names.
6. Click **View Errors** to review files that cannot be exported at all.
   The View Errors window opens (Figure 4.11).

Figure 4.11 View Errors

The window displays a list of files and data with the following information:

Experiment	Name of exported experiment.
Data Name	Particular file causing problem (DAT, template, etc.).
Data Type	Type of file.
Error	Reason the file cannot be exported.

7. Click **Close** in the View Error window.
8. Click **Export**.
   The Export Status window opens (Figure 4.12).

Export Status

This window (Figure 4.12) displays the status of the data export.

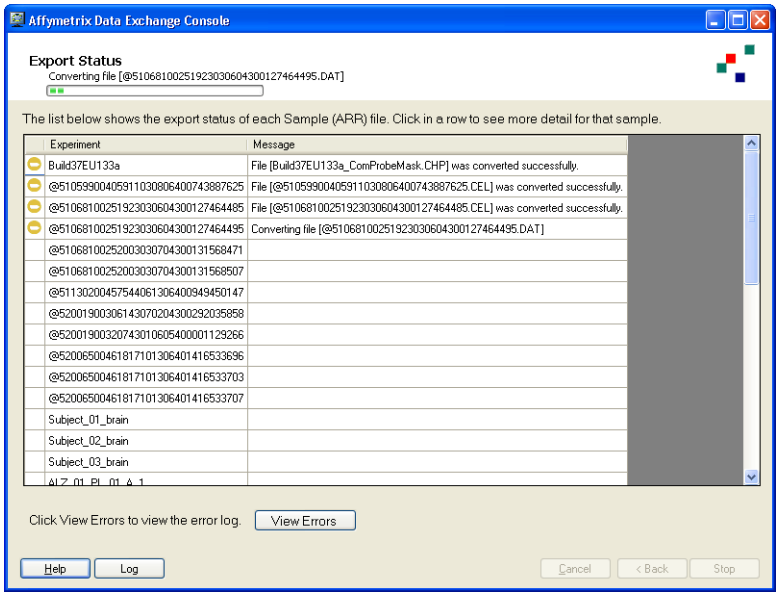

Figure 4.12 Export Status

The window displays the following columns:

Status Icon	Indicator of the status of the export: No conflicts or conflicts resolved. In progress. Error condition (File will not be exported).
Experiment	Name of exported experiment
Message	Particular file causing problem (DAT, template, etc.)

After the export is finished:

The DTT file is created.

- Click **View Errors** to view the error log; or
  Click **Start Again** to export more files; or
  Click **Finish** to close the DEC window.
  The exported data is visible in the AGCC Data folder.

Chapter 5

IMPACTS ON DATA

There are differences in the way GCOS organizes and stores data and the way AGCC does. These differences can cause impacts when you copy data from one to the other.

The impacts may vary, depending upon:

- The software source for the data being imported (MAS 5.x and the different versions of GCOS)
- The type of file being transferred (DAT, CEL, and CHP)
- The array type (expression, reseq, mapping, etc.)
- The type of algorithm used for processing the data
- Which DEC process is used (file to file or system to system)

These impacts are described briefly in this chapter of the DEC manual.

- *Impacts Caused by Migrating from GCOS to AGCC* on page 47
- *Impacts Caused by Migrating from AGCC to GCOS* on page 50

Impacts Caused by Migrating from GCOS to AGCC

The differences in file formats and data organization between GCOS and AGCC may cause impacts to the data being copied.

The particular impacts depend upon the type of data being transferred.

When you copy data from GCOS to AGCC, the following files and data may be impacted:

- *DAT Files* on page 48
- *CEL Files* on page 48
- *CHP Files* on page 49
- *Sample Data* on page 49
- *Experiment Data* on page 49
- *Templates* on page 50
- *Usersets* on page 50

NOTE: Problems that can be caused by exporting data using DTT from different versions of GCOS are described in *Impacts Caused by Migrating from AGCC to GCOS* on page 50.

DAT Files

Converting GCOS DAT files to Command Console format preserves the pixel intensity and both the global and subgrid coordinates. The Command Console DAT format stores the global grid and subgrid pixel coordinates as floating point values to support sub-pixel gridding. GCOS DAT files save the global grid pixel coordinates as integral values and subgrid pixel coordinates as floating point values.

The GCOS DAT files stores the cell coordinates of each subgrid. This information no longer exists in the Command Console DAT file. The subgrid cell coordinates are instead read from the SMD file in Command Console.

The Command Console DAT file preserves most of the GCOS DAT header information but not all. The following lists the GCOS DAT header data that has been removed or may be missing from the Command Console DAT header.

GCOS Data Removed from AGCC files:

- Scan Power
- Scan Temp
- Scan Speed
- Mean of the pixel intensities
- Standard deviation of the pixel intensities
- Average Offset
- Offset Stddev
- Offset samples
- Scan Comments
- Scan Operator
- Verify Grid
- Grid algorithm parameters
- Grid algorithm name
- Number of scans
- Any GCOS DAT data that is missing and is present in the GCOS system

The barcode has been added to the Command Console DAT file header.

CEL Files

Converting GCOS CEL files to Command Console format preserves the CEL intensity, error (stdev), number of pixels, and the mask and outlier flags.

The Command Console CEL file preserves most of the GCOS CEL header information including the global grid coordinates from the DAT file, the cell rows and cell columns, the probe array type and the algorithm parameters. The following table lists other header information that is not preserved in Command Console:

- OffsetX
- OffsetY
- Axis-invertX
- Axis-invertY
- swapXY
- Algorithm parameters not in GCOS CEL
- CEL Algorithm name

CHP Files

CHP files created with the new expression algorithms (e.g., PLIER, RMA, iterPLIER) and genotyping algorithms (e.g., BRLMM_P, Birdseed) cannot be loaded into GCOS and will not be included in the DTT file.

Sample Data

Converting GCOS sample data to Command Console format preserves the following data:

- sample name
- sample type
- sample project
- sample user
- sample date
- sample template name
- all sample template attribute names, types and values, with the following exceptions:
  - Integer is converted to floating point number
  - Time attribute is converted to text

In Command Console, the sample data is saved in the Sample (ARR) file.

Experiment Data

Converting GCOS experiment data to Command Console format preserves the following data:

- experiment name
- array barcode
- probe array type
- assay type
- userset name
- automation flag
- experiment date
- all experiment template attribute names, types and values, with the following exceptions:
  - Integer is converted to floating point number
  - Time attribute is converted to text

This information is saved in the Physical Array element of the Sample file in Command Console.

Most of the GCOS scan information is written to the Command Console DAT header. See the DAT file section for further details.

The GCOS hybridization information is preserved in the Command Console array audit log. The log contains:

- protocol name
- operator name
- date
- device name
- station ID
- station number
- module number
- number of stages
- each step in the protocol

Templates

Converting GCOS templates to Command Console format removes the distinction between a sample and an experiment template and removes the notion of an 'active' template. The template name and attribute names, types, values, controlled value lists, and required flags are preserved in template files.

The imported template files are available for the user in Command Console.

Usersets

Converting GCOS usersets to Command Console results in one or more PARAM files per userset; one for each probe array type defined in the userset. The format of the PARAM file name includes the userset name and probe array type. The name, type and value of each parameter is preserved.

Impacts Caused by Migrating from AGCC to GCOS

You can export data from AGCC in DTT format and import it into different versions of GCOS. There are some issues that arise when you do that, described in this section.

The AGCC data is exported into DTT files using GCOS 1.4 formats. When the DTT file is imported into GCOS, it will be converted to earlier GCOS formats (GCOS 1.x) if necessary.

The GCOS versions are backward compatible, but not forward compatible. That is, you can view a GCOS 1.1 file in GCOS 1.4, but you may not be able to use a GCOS 1.4 file in GCOS 1.1.

GCOS File Format Changes

The file formats used in different versions of GCOS have changed to accommodate new features and to improve performance. These file format changes may cause data impacts when importing or export data from GCOS and AGCC.

The file formats that may change are summarized in the table below.

Table 5.1 Software versions and data file versions

	MAS 5.x	GCOS 1.0, 1.1.x	GCOS 1.2	GCOS 1.3	GCOS 1.4*
DAT File	DAT v1	DAT v1	DAT v1	DAT v2 (Subgrids added)	DATv2 (Subgrids added)
CEL file	CEL v1 (ASCII format)	CEL v2 (Binary Version)	CEL v2 (Binary Version)	CEL v2 (Binary Version)	CEL v2 (Binary Version)
CHP (Expression)	CHP v1	CHP v2	CHP v3 (XDA v1)	CHP v3 (XDA v1)	
CHP (Universal)	CHP v1	CHP v2	CHP v3 (XDA v1)	CHP v3 (XDA v1)	
CHP (Mapping)		CHP v2	CHP v3 (XDA v1)	CHP v3 (XDA v1)	CHP v3 (XDA v1)
CHP (Resequencing)		CHP v2	CHP v3 (XDA v1)	CHP v3 (XDA v1)	CHP v4 (XDA v2)
RPT	RPT v1				

** Format used by DEC for output*